

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:31:29 ; Search time 140.273 Seconds
(without alignments)
408.065 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPHKHSHEQHPGHHPAH.....PSFPLPHHKHPLKPDNQPPF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	966	100.0	148	5	ABB79807	Abb79807 Rabbit hi
2	966	100.0	148	5	ABB79806	Abb79806 Human his
3	966	100.0	148	8	ADH10411	Adh10411 Human HPR
4	966	100.0	525	5	ABB79804	Abb79804 Human his
5	966	100.0	525	8	ADH10409	Adh10409 Human HPR
6	949	98.2	525	8	ADE76897	Ade76897 Human pro
7	949	98.2	525	8	ADL12335	Adl12335 Human ste
8	543	56.2	526	5	ABB79805	Abb79805 Rabbit hi
9	543	56.2	526	8	ADH10410	Adh10410 Rabbit HP
10	302	31.3	75	5	ABB07123	Abb07123 HRGP thro
11	280.5	29.0	101	8	ADH10412	Adh10412 Rabbit HP
12	254.5	26.3	245	4	ABG28019	Abg28019 Novel hum
13	251.5	26.0	330	4	ABG25331	Abg25331 Novel hum
14	248.5	25.7	406	4	ABG27250	Abg27250 Novel hum
15	242.5	25.1	233	4	ABG19767	Abg19767 Novel hum
16	242.5	25.1	378	8	ABO60258	Abo60258 Human gen
17	241.5	25.0	351	2	AAR24393	Aar24393 Sequence
18	240	24.8	491	4	ABG09949	Abg09949 Novel hum
19	230	23.8	183	4	ABG25337	Abg25337 Novel hum
20	225.5	23.3	598	4	ABG23595	Abg23595 Novel hum
21	221	22.9	292	4	AAM18326	Aam18326 Peptide #
22	221	22.9	292	4	ABB37360	Abb37360 Peptide #
23	221	22.9	292	4	AAM30814	Aam30814 Peptide #
24	221	22.9	292	4	ABB32108	Abb32108 Peptide #
25	221	22.9	292	4	ABB22646	Abb22646 Protein #

26	221	22.9	292	4	AAM70490	Aam70490 Human bon
27	221	22.9	292	4	AAM58050	Aam58050 Human bra
28	221	22.9	292	4	ABG52171	Abg52171 Human liv
29	221	22.9	292	4	AAM05933	Aam05933 Peptide #
30	221	22.9	292	5	ABG40129	Abg40129 Human pep
31	219.5	22.7	227	4	ABG14399	Abg14399 Novel hum
32	215	22.3	124	4	ABG12879	Abg12879 Novel hum
33	210	21.7	82	4	AAM13580	Aam13580 Peptide #
34	210	21.7	82	4	ABB32509	Abb32509 Peptide #
35	210	21.7	82	4	AAM25978	Aam25978 Peptide #
36	210	21.7	82	4	ABB27363	Abb27363 Human pep
37	210	21.7	82	4	ABB18015	Abb18015 Protein #
38	210	21.7	82	4	AAM65721	Aam65721 Human bon
39	210	21.7	82	4	AAM53343	Aam53343 Human bra
40	210	21.7	82	4	ABG47362	Abg47362 Human liv
41	210	21.7	82	4	AAM01333	Aam01333 Peptide #
42	210	21.7	82	5	ABG35350	Abg35350 Human pep
43	209.5	21.7	324	4	ABG20509	Abg20509 Novel hum
44	209	21.6	304	4	ABG08412	Abg08412 Novel hum
45	208	21.5	86	5	ADK36683	Adk36683 Novel hum

ALIGNMENTS

RESULT 1
ABB79807
ID ABB79807 standard; protein; 148 AA.
XX
AC ABB79807;
XX
DT 25-NOV-2002 (first entry)
XX
DE Rabbit histidine proline rich glycoprotein His/Pro-rich domain.
XX
KW Histidine proline rich glycoprotein; HPGP; rabbit; antiangiogenic;
KW cyostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
XX
OS Oryctolagus cuniculus.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
PS Claim 1; Page 13; 82pp; English.
XX
CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of rabbit anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79805). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain (see ABB79806) of human
CC HPRG; an H/P domain of rabbit HPRG; a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide

CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX SQ Sequence 148 AA;

Query Match 100.0%; Score 966; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 5.2e-80;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPGHHPHCH 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
1 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHCH 60

QY 61 DFQDYGPCDPPPHNQHCCHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
61 DFQDYGPCDPPPHNQHCCHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120

QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
Db |||||||||||||||||||||||||||||||||||||||||||||||||
121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148

RESULT 2
ABB79806
ID ABB79806 standard; protein; 148 AA.
XX
AC ABB79806;

XX 25-NOV-2002 (first entry)
XX Human histidine proline rich glycoprotein His/Pro-rich domain.
DE
XX Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cytosstatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian.

XX Homo sapiens.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX

PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 13; 82pp; English.

XX The present sequence is the protein sequence of the histidine-proline-
rich (H/P) domain of human anti-angiogenic histidine proline rich
glycoprotein (HPRG, see also ABB79804). Claimed anti-angiogenic
polypeptides or peptides comprise: the H/P domain of human HPRG; an H/P
domain of rabbit HPRG (see ABB12345); a variant of these that is capable
of inhibiting angiogenesis, endothelial cell proliferation or endothelial
tube formation in vitro or in vivo; or a pentapeptide having the generic
sequence given in ABB79808. Also claimed are: chemically synthesised or
recombinantly produced peptide multimers; a diagnostically or peptide
therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
multimer; a diagnostically useful HPRG-related composition, comprising
the diagnostically labeled polypeptide, peptide or peptide multimer and a
carrier; an antibody specific for an epitope of HPRG that is present in
the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
of the domains in a way which inhibits the anti-angiogenic activity of
HPRG or the domain, or an antigen-binding fragment of the antibody; a
method for inhibiting cell migration, cell invasion, cell proliferation
or angiogenesis, or for inducing apoptosis; a method for treating a
subject having a disease or condition associated with undesired cell
migration, invasion, proliferation, or angiogenesis; a method of
stimulating or inhibiting angiogenesis in a subject; a method of
detecting the presence of HPRG or its cleavage product or its peptide in
a biological sample; isolated nucleic acids encoding the polypeptide,
peptide or peptide multimer; an expression vector; transformed or
transfected cells; a method of providing to a cell, tissue or organ an
angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
pentapeptide, or the peptide multimer; an affinity ligand useful for
binding to, or isolating, an HPRG-binding molecule or cells expressing
the binding molecule, comprising the polypeptide, peptide or peptide
multimer, immobilised to a solid support or carrier; and a method of
isolating HPRG-binding molecule, or isolating or enriching cells
expressing HPRG-binding site or receptor, from a complex mixture. The
compositions and methods are useful in diagnosing or treating a disease
or condition associated with undesired cell migration, invasion,
proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
fractures. HPRG is especially useful in inhibiting the growth of primary
tumours or metastases, and may also be used in treating neurodegenerative
diseases like Alzheimer's or Parkinson's disease. The antibodies are
stimulators of angiogenesis and are useful for promoting
neovascularization in pertinent disease states, and in various
immunoassays

XX SQ Sequence 148 AA;

Query Match 100.0%; Score 966; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 5.2e-80;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHCH 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
1 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHCH 60

QY 61 DFQDYGPCDPPPHNQHCCHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||
61 DFQDYGPCDPPPHNQHCCHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120

CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
CC useful HPRG-related composition, comprising the diagnostically labeled
CC polypeptide, peptide or peptide multimer and a carrier; an antibody
CC specific for an epitope of HPRG that is present in the H/P domain of
CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,
CC or an antigen-binding fragment of the antibody; a method for inhibiting
CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
CC inducing apoptosis; a method for treating a subject having a disease or
CC condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis; a method of stimulating or inhibiting
CC angiogenesis in a subject; a method of detecting the presence of HPRG or
CC its cleavage product or its peptide in a biological sample; isolated
CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an
CC expression vector; transformed or transfected cells; a method of
CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of
CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;
CC an affinity ligand useful for binding to, or isolating, an HPRG-binding
CC molecule or cells expressing the binding molecule, comprising the
CC polypeptide, peptide or peptide multimer, immobilised to a solid support
CC or carrier; and a method of isolating HPRG-binding molecule, or isolating
CC or enriching cells expressing HPRG-binding site or receptor, from a
CC complex mixture. The compositions and methods are useful in diagnosing or
CC treating a disease or condition associated with undesired cell migration,
CC invasion, proliferation, or angiogenesis, such as cancer,
CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,
CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in
CC inhibiting the growth of primary tumours or metastases, and may also be
CC used in treating neurodegenerative diseases like Alzheimer's disease,
CC Parkinson's disease, and amyotrophic lateral sclerosis. The antibodies
CC are stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX Sequence 525 AA;

Query Match 100.0%; Score 966; DB 5; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.9e-79;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHPHGHHPHCH 60
Db |||||||
350 HPHKHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 120
Db |||||||
410 DFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 469

QY 121 LPLPEANFPSFPLPHKHPLKPDNQPPF 148
Db |||||||
470 LPLPEANFPSFPLPHKHPLKPDNQPPF 497

RESULT 5
ADH10409
ID ADH10409 standard; protein; 525 AA.
XX
AC ADH10409;

XX 11-MAR-2004 (first entry)

XX Human HPRG protein sequence.

XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiac; vasotropic; vulnery;

XX angiogenesis; gene therapy; human.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 276. .321

FT Domain /note = pro-rich domain
FT 350. .497
FT /note = His-Pro rich domain

XX WO2003077872-A2.

XX 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

XX 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

PI Mcrae K, Donate F, Juarez J, Mazar AP;

XX WPI; 2004-090604/09.

XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.

PS Claim 6; SEQ ID NO 21; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a human HPRG
CC protein sequence.

XX Sequence 525 AA;

Query Match 100.0%; Score 966; DB 8; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.9e-79;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHPHGHHPHCH 60
Db |||||||
350 HPHKHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 120
Db |||||||
410 DFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 469

QY 121 LPLPEANFPSFPLPHKHPLKPDNQPPF 148
Db |||||||
470 LPLPEANFPSFPLPHKHPLKPDNQPPF 497

RESULT 6
ADE76897

ID AD E76897 standard; protein; 525 AA.
XX AC ADE76897;
XX DT 29-JAN-2004 (first entry)
XX DE Human protein expressed in a liver disorder #18.
XX KW human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
OS Homo sapiens.
XX PN US2003108871-A1.
XX PD 12-JUN-2003.
XX PF 30-JUL-2001; 2001US-00919039.
XX PR 28-JUL-2000; 2000US-0222113P.
XX PA (KASE/) KASER M R.
XX PI Kaser MR;
XX DR WPI; 2004-031227/03.
DR N-PSDB; ADE76896.
XX Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX Claim 1; SEQ ID NO 62; 41pp; English.
PS
XX The invention relates to a composition comprising several cDNAs that are
CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.
XX
SQ Sequence 525 AA;
Query Match 98.2%; Score 949; DB 8; Length 525;
Best Local Similarity 98.6%; Pred. No. 6.6e-78;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 HPHKHSHEQHPHGHPHAHPHEHDTHQHPHGHPHGHPHGHPHGHPHGHPHGHPCH 60
Db 350 HPHKHSHEQHPHGHPHAHPHEHDTHQHPHGHPHGHPHGHPHGHPHGHPHGHPCH 409
QY 61 DFQDYGPCDPPPHNQCHCHGPPPGHLRRRCGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db 410 DFQDYGPCDPPPHNQCHCHGPPPGHLRRRCGKGRPFHCRQIGSVYRLPPLRKGEV 469
QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
Db 470 LPLPEANFSPFLPHHKHPLKPDIOQPPF 497

RESULT 7
ADL12335
ID ADL12335 standard; protein; 525 AA.
XX AC ADL12335;
XX DT 06-MAY-2004 (first entry)
XX DE Human steroid-induced C3A liver cell protein #11.
XX KW Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX OS Homo sapiens.
XX PN US6673549-B1.
XX PD 06-JAN-2004.
XX PF 12-OCT-2001; 2001US-00976594.
XX PR 12-OCT-2000; 2000US-0240409P.
XX PA (INCY-) INCYTE CORP.
XX PI Furness LM, Buchbinder JL;
XX DR WPI; 2004-068610/07.
XX Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
PS Disclosure; SEQ ID NO 64; 141pp; English.
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human protein which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 525 AA;
Query Match 98.2%; Score 949; DB 8; Length 525;
Best Local Similarity 98.6%; Pred. No. 6.6e-78;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 HPHKHSHEQHPHGHPHAHPHEHDTHQHPHGHPHGHPHGHPHGHPHGHPHGHPCH 60
Db 350 HPHKHSHEQHPHGHPHAHPHEHDTHQHPHGHPHGHPHGHPHGHPHGHPHGHPCH 409
QY 61 DFQDYGPCDPPPHNQCHCHGPPPGHLRRRCGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db 410 DFQDYGPCDPPPHNQCHCHGPPPGHLRRRCGKGRPFHCRQIGSVYRLPPLRKGEV 469
QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148

Db 470 LPLPEANFSPFLPHHKHPLKPDIQPPF 497

RESULT 8

ABB79805

ID ABB79805 standard; protein; 526 AA.

XX

AC ABB79805;

XX

DT 25-NOV-2002 (first entry)

XX

DE Rabbit histidine proline rich glycoprotein.

XX

KW Histidine proline rich glycoprotein; HPRG; rabbit; antiangiogenic;

KW cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;

KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;

KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.

XX

OS Oryctolagus cuniculus.

XX

FH Key Location/Qualifiers

FT Domain 251..296

FT /note= "Proline-rich domain"

FT Domain 321..421

FT /note= "Histidine-proline-rich domain, region also

FT specifically claimed in Claim 1"

XX

PN WO200264621-A2.

XX

PD 22-AUG-2002.

XX

PF 14-FEB-2002; 2002WO-US004336.

XX

PR 14-FEB-2001; 2001US-0268370P.

XX

PA (ATTE-) ATTENUON LLC.

XX

PI Donate F, Harris S, Plunkett ML, Mazar AP;

XX

DR WPI; 2002-666989/71.

DR P-PSDB; ABN84911.

XX

PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for

PT diagnosing or treating diseases associated with undesired cell migration,

PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX

PS Claim 1; Page 12-13; 82pp; English.

XX

CC The present sequence is the protein sequence of rabbit histidine proline

CC rich glycoprotein (HPRG), a proteinaceous chelator that can be used to

CC inhibit angiogenesis and treat cancer. Claimed anti-angiogenic

CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)

CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see

CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,

CC endothelial cell proliferation or endothelial tube formation in vitro or

CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.

CC Also claimed are: chemically synthesised or recombinantly produced

CC peptide multimers; a diagnostically or therapeutically labeled anti-

CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically

CC useful HPRG-related composition, comprising the diagnostically labeled

CC polypeptide, peptide or peptide multimer and a carrier; an antibody

CC specific for an epitope of HPRG that is present in the H/P domain of

CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in

CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,

CC or an antigen-binding fragment of the antibody; a method for inhibiting

CC cell migration, cell invasion, cell proliferation or angiogenesis, or for

CC inducing apoptosis; a method for treating a subject having a disease or

CC condition associated with undesired cell migration, invasion,

CC proliferation, or angiogenesis; a method of stimulating or inhibiting

CC angiogenesis in a subject; a method of detecting the presence of HPRG or

CC its cleavage product or its peptide in a biological sample; isolated

CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an

CC expression vector; transformed or transfected cells; a method of

CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of

CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;

CC an affinity ligand useful for binding to, or isolating, an HPRG-binding

CC molecule or cells expressing the binding molecule, comprising the

CC polypeptide, peptide or peptide multimer, immobilised to a solid support

CC or carrier; and a method of isolating HPRG-binding molecule, or isolating

CC or enriching cells expressing HPRG-binding site or receptor, from a

CC complex mixture. The compositions and methods are useful in diagnosing or

CC treating a disease or condition associated with undesired cell migration,

CC invasion, proliferation, or angiogenesis, such as cancer,

CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,

CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in

CC inhibiting the growth of primary tumours or metastases, and may also be

CC used in treating neurodegenerative diseases like Alzheimer's disease,

CC Parkinson's disease and amyotrophic lateral sclerosis. The antibodies are

CC stimulators of angiogenesis and are useful for promoting

CC neovascularization in pertinent disease states, and in various

CC immunoassays

XX

SQ Sequence 526 AA;

Query Match 56.2%; Score 543; DB 5; Length 526;

Best Local Similarity 57.6%; Pred. No. 4.9e-41;

Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

QY 2 PHKHSHSQHPHGHHPHA-----HH-----PHEHDTHQHPHGHHGHHP 42

Db 330 PHGHHPHGPPPHGHHPHGPPPHGHPHPGPPRHRPPHGGPPPHGHPHPGHPGPP 389

QY 43 HGHHPHGHHPHGHHPHCHDFQDYGPCDPPHNOG-----HCCHGHGPPPGHLRRCPGKGP 98

Db 390 HGHPHPHGPPPHGHPHPHGHHGFDHGPCDPPSHKGGPQLH-QHAMGPPPKHGKRGKGKH 448

QY 99 RPFHCRQIGSVYRLPPLRKGEVLPPEANFSPFLPHHKHPLKPDNQPPF 148

Db 449 FPFHRRIGSVYQLPPLQKGEVLPPEANFPQLLRNHTHPLKPEIQPPF 498

RESULT 9

ADH10410

ID ADH10410 standard; protein; 526 AA.

XX ADH10410;

AC ADH10410;

DT 11-MAR-2004 (first entry)

XX Rabbit HPRG protein sequence.

DE

XX Tpm; tropomyosin; antiangiogenic receptor;

KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;

KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;

KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;

XX angiogenesis; gene therapy; rabbit.

OS Oryctolagus cuniculus.

XX

FH Key Location/Qualifiers

FT Domain 251..296

FT /note = pro-rich domain

FT Domain 321..421

FT /note = His-Pro rich domain

XX WO2003077872-A2.

PN

XX 25-SEP-2003.

PD

XX 17-MAR-2003; 2003WO-US008060.

PF

XX 15-MAR-2002; 2002US-0364047P.

PR

XX (ATTE-) ATTENUON LLC.

XX Mccrae K, Donate F, Juarez J, Mazar AP;

PI

DR WPI; 2004-090604/09.

XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for

PT inhibiting endothelial cell migration, invasion, proliferation or

PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or

PT cancer.

XX Claim 6; SEQ ID NO 22; 117pp; English.

PS

XX The invention relates to an isolated tropomyosin (Tpm)-related

CC antiangiogenic receptor polypeptide or peptide, which is a fragment

CC of a full-length native Tpm protein expressed on the surface of

CC endothelial cells, or a variant of the fragment. It has a molecular mass

CC of about 17 kDa and corresponds in its sequence to, or is a variant of,

CC an internal fragment of a native Tpm isoform which is a binding site for

CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor

CC polypeptide, peptide or variant has substantially the same biochemical

CC activity of binding to the antiangiogenic polypeptide agents, as does the

CC native Tpm internal fragment. The antiangiogenic polypeptide agent that

CC binds to the isolated polypeptide or peptide is human histidine-proline

CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic

CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain

CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,

CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5

CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,

CC antibodies and compositions are useful for inhibiting endothelial cell

CC migration, invasion, proliferation or angiogenesis, for inducing

CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic

CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,

CC psoriasis, or scleroderma. The antibody may be also used for detecting

CC the presence of a Tpm polypeptide or peptide in a biological sample, for

CC promoting wound healing, or for treating diseases or conditions in which

CC increased angiogenesis is desired, e.g. coronary artery disease or

CC peripheral artery disease. The present sequence represents a rabbit HPRG

CC protein sequence.

XX

SQ Sequence 526 AA;

Query Match 56.2%; Score 543; DB 8; Length 526;

Best Local Similarity 57.6%; Pred. No. 4.9e-41;

Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

QY 2 PHKHSHEQHPGHHPHA-----HH-----PHEHDTHRQHPGHHPGHHP 42

Db 330 PHGHPHGPpppHGHPGHPPPHGHPGHPPPHGHPGHPPPHGHPGHPP 389

QY 43 HGHHPGHHPGHHPCHDFQDYGPCDPPPHNQG---HCCHGHPGHRRRGGKGP 98

Db 390 HGHPPHGPpppHGHPGHGFHDHGCDPPSHKEGQDLH-QHAMGPPPKHKGKGGKH 448

QY 99 RPFHCRQIGSVYRLPLRKGEVLPPLPEANFSPFPLPHKHPLKPDNQPPF 148

Db 449 FPFHWRRIQSVYQLPPLQKGEVLPPLPEANFPQLLLRNHTHPLKPEIQPPF 498

RESULT 10

ABB07123

ID ABB07123 standard; protein; 75 AA.

XX

AC ABB07123;

XX

DT 13-MAR-2002 (first entry)

XX

DE HPRG thrombospondin (TSP)-binding domain fragment (residues 443-517).

XX

KW Thrombospondin; TSP-1; histidine-rich glycoprotein; HPRG; cytostatic;

KW cardiovascular; protein therapy; gene therapy; antisense inhibition;

KW ribozyme inhibition; antibody inhibition; angiogenesis; human; cancer;

KW TSP.

XX

OS Homo sapiens.

XX

PN WPI; 2004-090604-A1.

XX 15-NOV-2001.

XX 05-DEC-2000; 2000US-00730379.

XX 06-DEC-1999; 99US-0169205P.

XX (SIMA/) SIMANTOV R.

PA (SILV/) SILVERSTEIN R L.

XX Simantov R, Silverstein RL;

XX WPI; 2002-017220/02.

DR Thrombospondin-binding molecules and fragments comprising regions of the

PT TSP-1 binding domains of Histidine-Rich Glycoproteins, useful for

PT treating cancers and cardiovascular diseases.

XX Example; Fig 1; 22pp; English.

PS The invention provides thrombospondin (TSP)-binding molecules and

CC fragments comprising regions of the TSP-1 binding domains of Histidine-

CC Rich Glycoproteins (HRGP). The proteins are used for the regulation of

CC cellular growth and proliferation, such as in the accretion of new blood

CC vessels (angiogenesis), particularly for treatment of cardiovascular

CC disease. They may be used for the reduction of angiogenesis, including

CC growth suppression and arrest, and apoptosis in normal development, for

CC example in embryogenesis, and in a wide range of disorders and diseases,

CC including those involving tumors, malignancies, neoplastic and other

CC pathological conditions and homeostatic imbalances in the control of

CC growth and development, in particular anal cancer, bladder cancer, small

CC cell lung cancer, non-small cell lung cancer, bone cancer, brain cancer,

CC breast cancer, cervical cancer, chondrosarcoma, clear cell adenocarcinoma

CC (DES), colorectal cancer, endometrial cancer, oesophageal cancer, cancers

CC of the eye, eyelid, Kaposi's sarcoma, kidney cancer, cancer of the

CC larynx, leiomyosarcoma, leukemia, liver cancer, lung cancer, lymphoma,

CC melanoma, mesothelioma, oral cancer, ovarian cancer, pancreatic cancer,

CC prostate cancer, skin cancer, squamous cell cancer, stomach cancer,

CC testicular cancer, thyroid cancer, hepatoma, neuroendocrine cancer,

CC liposarcoma, head and neck cancer and a cholangiocarcinoma. The present

CC sequence represents the TSP-binding motif of HRGP

XX

SQ Sequence 75 AA;

Query Match 31.3%; Score 302; DB 5; Length 75;

Best Local Similarity 100.0%; Pred. No. 5.3e-20;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 GPRPFHCRQIGSVYRLPPLRKGEVLPPLPEANFSPFPLPHKHPLKPDNQPPF 148

Db 1 GPRPFHCRQIGSVYRLPPLRKGEVLPPLPEANFSPFPLPHKHPLKPDNQPPF 52

RESULT 11

ADH10412

ID ADH10412 standard; protein; 101 AA.

XX

AC ADH10412;

XX

DT 11-MAR-2004 (first entry)

XX

DE Rabbit HPRG protein H/P rich domain.

XX

KW Tpm; tropomyosin; antiangiogenic receptor;

KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;

KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;

KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;

KW angiogenesis; gene therapy; rabbit.

XX

OS Oryctolagus cuniculus.

XX

PN WO2003077872-A2.

XX PD 25-SEP-2003.
XX PF 17-MAR-2003; 2003WO-US008060.
XX PR 15-MAR-2002; 2002US-0364047P.
XX PA (ATTE-) ATTENUON LLC.
XX PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX PR WPI; 2004-090604/09.
XX PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX PS Claim 6; SEQ ID NO 24; 117pp; English.
XX CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC an antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX SQ Sequence 101 AA;

Query Match 29.0%; Score 280.5; DB 8; Length 101;
Best Local Similarity 51.6%; Pred. No. 6.4e-18;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;
QY 2 PHKHSHEQHPHGHHPHA-----HH-----PHEHDTHRQHPHGHHPHGHHP 42
Db 10 PHGHHPHGPPPHGHHPHGPPPHGHPHGRHPPHGGPPPHGHPHGRHPPHGGPPPHGPP 69
QY 43 HGHHPHGHHPHGHHPHCHDFQDYGPCDPPPH 73
Db 70 HGHPPHGGPPPHGHPHGHGFDHGGCDPPSH 100

RESULT 12
ABG28019
ID ABG28019 standard; protein; 245 AA.
XX AC ABG28019;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #28010.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS92206.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 20; SEQ ID NO 58378; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 245 AA;

Query Match 26.3%; Score 254.5; DB 4; Length 245;
Best Local Similarity 48.2%; Pred. No. 3.6e-15;
Matches 41; Conservative 2; Mismatches 39; Indels 3; Gaps 1;
QY 1 HPHKHSHEQHPHGHHPHAHPHHEHDTHRQHPHGHHPHGHHPHGHHPHGH---HP 57
Db 25 HYHHHHHHHHYHRRHHHPHHPHHPHHPHYHTTTHHHHHHHHHYHQHHPHHHHHHHLLHH 84
QY 58 HCHDFQDYGPCDPPPHNQHCCHGH 82
Db 85 HHHHHHQHHPHHPHHPHHHHHTHHHH 109

RESULT 13
ABG25331
ID ABG25331 standard; protein; 330 AA.
XX AC ABG25331;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #25322.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS89518.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID NO 55690; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 330 AA;

Query Match 26.0%; Score 251.5; DB 4; Length 330;

Best Local Similarity 39.3%; Pred. No. 9.1e-15;

Matches 46; Conservative 5; Mismatches 43; Indels 23; Gaps 3;

QY 1 HPHKHS---HEQHPHGHHPHAHHPHEHDTHQHPHGHHPHGHHPGHHPGHHPGHHP 57

Db 205 HPHHHHHYHHHHHHHHYHPYPHHYHHHHHHHHHPHHHHHHHHYHHHPHHHHHP 264

QY 58 HCHDFQDYGCDPPPHNQHCCHGHPGPHLRR-----RGPGKGP 98

Db 265 HHH----HRPHHHRHQHPHHYHHHPCHHRLHHHHHHHHYHHTSSVQPGSQP 317

RESULT 14

ABG27250

ID ABG27250 standard; protein; 406 AA.

XX

AC ABG27250;

XX

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #27241.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS91437.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID NO 57609; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 406 AA;

Query Match 25.7%; Score 248.5; DB 4; Length 406;

Best Local Similarity 32.6%; Pred. No. 2.1e-14;

Matches 45; Conservative 3; Mismatches 41; Indels 49; Gaps 3;

QY 1 HPHKHSHEQHPHGHHPHAHHPHEHDTHQHPHGHHPHGHHPGHHPGHHPGHCH 60

Db 206 HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCH 265

QY 61 DFQDYGCDPPPHNQHCCHGHPGPHLRRRGPGKGPFPFHCRCQIGSVYRLPPLRKGEV 120

Db 266 HH-----HHHHHCHHHHHCHHHHR-----HCH----- 290

QY 121 LPLPEANFFSFPLPHKH 138

Db 291 -----HHRH 294

RESULT 15

ABG19767
ID ABG19767 standard; protein; 233 AA.
XX AC ABG19767;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19758.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
DR N-PSDB; AAS83954.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 20; SEQ ID NO 50126; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 233 AA;

Query Match 25.1%; Score 242.5; DB 4; Length 233;
Best Local Similarity 35.0%; Pred. No. 4.2e-14;
Matches 57; Conservative 8; Mismatches 71; Indels 27; Gaps 7;
QY 1 HPHKHSHEQ-HPHGHPHAHPHEH-----DTHRQHPGHHPGHG----HPHGHPHG 49
DB 52 HPHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 111
QY 50 HHPGHHPHCHFDQYGPCDPPPHNQHCCHGCHGPPPGHLRRRGPGKGRPFCHCRQIGSV 109
DB 112 HYSHHHHHPHHHXRHDHYAHHHHHQHCHCHCH-----HPRHGH-----HYHYHHHGTV 160

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:38:35 ; Search time 36.8514 Seconds
(without alignments)
299.800 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPHKHSHEQHPHGHPHAH.....PSFPLPHHKHPLKPDNQPPF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTus_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	949	98.2	525	4	US-09-976-594-64 Sequence 64, Appl
2	949	98.2	525	4	US-09-919-039-62 Sequence 62, Appl
3	185	19.2	26	3	US-08-785-636-1 Sequence 1, Appli
4	185	19.2	26	3	US-09-095-407-2 Sequence 2, Appli
5	185	19.2	26	4	US-09-742-494-2 Sequence 2, Appli
6	185	19.2	26	4	US-09-809-657-2 Sequence 2, Appli
7	184.5	19.1	167	3	US-09-507-323B-13 Sequence 13, Appl
8	180	18.6	163	4	US-09-902-540-13395 Sequence 13395, A
9	172.5	17.9	400	4	US-09-543-681A-6151 Sequence 6151, Ap
10	162.5	16.8	125	4	US-09-248-796A-24231 Sequence 24231, A
11	161.5	16.7	225	4	US-09-252-991A-28819 Sequence 28819, A
12	158.5	16.4	218	4	US-09-252-991A-25291 Sequence 25291, A
13	158	16.4	955	4	US-09-949-016-8369 Sequence 8369, Ap
14	156.5	16.2	971	4	US-09-248-796A-19531 Sequence 19531, A
15	155.5	16.1	865	3	US-09-281-766-19 Sequence 19, Appl
16	155.5	16.1	865	4	US-09-612-858-19 Sequence 19, Appl
17	155.5	16.1	865	4	US-09-957-995A-19 Sequence 19, Appl
18	148.5	15.4	89	3	US-09-507-323B-12 Sequence 12, Appl
19	148	15.3	130	4	US-09-270-767-39206 Sequence 39206, A
20	148	15.3	130	4	US-09-270-767-54423 Sequence 54423, A
21	147	15.2	339	4	US-09-328-352-6551 Sequence 6551, Ap
22	145	15.0	376	4	US-09-270-767-32753 Sequence 32753, A
23	145	15.0	376	4	US-09-270-767-47970 Sequence 47970, A
24	144.5	15.0	60	1	US-08-255-457-1 Sequence 1, Appli
25	144.5	15.0	60	2	US-09-115-032-1 Sequence 1, Appli
26	144.5	15.0	60	5	PCT-US95-05772-1 Sequence 1, Appli
27	140	14.5	189	1	US-08-152-922A-6 Sequence 6, Appli

28	140	14.5	309	1	US-08-161-406-2	Sequence 2, Appli
29	136.5	14.1	1319	4	US-09-538-092-1291	Sequence 1291, Ap
30	135.5	14.0	297	4	US-09-489-039A-12802	Sequence 12802, A
31	134	13.9	398	3	US-09-461-474-17	Sequence 17, Appl
32	132	13.7	203	4	US-09-270-767-35326	Sequence 35326, A
33	132	13.7	203	4	US-09-270-767-50543	Sequence 50543, A
34	132	13.7	313	3	US-08-686-528A-3	Sequence 3, Appli
35	132	13.7	313	3	US-09-456-287-3	Sequence 3, Appli
36	132	13.7	337	3	US-08-686-528A-2	Sequence 2, Appli
37	132	13.7	337	3	US-09-456-287-2	Sequence 2, Appli
38	131.5	13.6	1343	4	US-09-949-016-10641	Sequence 10641, A
39	130	13.5	413	4	US-09-252-991A-28004	Sequence 28004, A
40	129	13.4	693	4	US-09-949-016-7806	Sequence 7806, Ap
41	128.5	13.3	474	3	US-09-461-474-10	Sequence 10, Appl
42	128.5	13.3	613	4	US-09-252-991A-24982	Sequence 24982, A
43	128.5	13.3	754	4	US-09-252-991A-22309	Sequence 22309, A
44	127.5	13.2	623	4	US-09-949-016-6530	Sequence 6530, Ap
45	127	13.1	264	4	US-09-270-767-44345	Sequence 44345, A

ALIGNMENTS

RESULT 1
US-09-976-594-64
; Sequence 64, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 085596CD1
US-09-976-594-64

Query Match	98.2%;	Score 949;	DB 4;	Length 525;
Best Local Similarity	98.6%;	Pred. No. 2.1e-85;		
Matches 146;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	HPHKHSHEQHPHGHPHAHPHEHDT	THRQHPHGHPHGHPHGHPHGHPCH	60
Db	350	HPHKHSHEQHPHGHPHAHPHEHDT	THRQHPHGHPHGHPHGHPHGHPCH	409
QY	61	DFQDYGPCDPPPHNQHCCHGHPGPP	PGHLRRRPGKGRPFHCRQIGSVYRLP	120
Db	410	DFQDYGPCDPPPHNQHCCHGHPGPP	PGHLRRRPGKGRPFHCRQIGSVYRLP	469
QY	121	LPLPEANFPSPFLPHHKHPLKPDNQPP	F 148	
Db	470	LPLPEANFPSPFLPHHKHPLKPDNQPP	F 497	

RESULT 2
US-09-919-039-62
; Sequence 62, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039

```

; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 085596CD1
US-09-919-039-62

Query Match          98.2%; Score 949; DB 4; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.1e-85;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 HPHKHSHEQHPGHHPHAHPHEHDTHQHPHGHHPHGHHPHGHHPHGHHPHCH 60
        |||||||
Db      350 HPHKHSHEQHPGHHPHAHPHEHDTHQHPHGHHPHGHHPHGHHPHGHHPHCH 409

Qy      61 DFQDYGPCDPPPHNQHCCHGPPPGHLLRRRPGKGRPFHCTQIGSVYRLPPLRKGEV 120
        |||||||
Db      410 DFQDYGPCDPPPHNQHCCHGPPPGHLLRRRPGKGRPFHCTQIGSVYRLPPLRKGEV 469

Qy      121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
        |||||||
Db      470 LPLPEANFSPFLPHHKHPLKPDIQPPF 497

RESULT 3
US-08-785-636-1
; Sequence 1, Application US/08785636
; Patent No. 6027942
; GENERAL INFORMATION:
; APPLICANT: Yip, Tai-Tung
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Method and Apparatus for Desorption and Ionization of
; TITLE OF INVENTION: Analytes
; FILE REFERENCE: Hutchens
; CURRENT APPLICATION NUMBER: US/08/785,636
; CURRENT FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/068,896
; EARLIER FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: peptide

US-08-785-636-1

Query Match          19.2%; Score 185; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      34 GHHPGHHPGHHPGHHPGHHPGHHPH 58
        |||||||
Db      1 GHHPGHHPGHHPGHHPGHHPGHHPH 25

RESULT 4
US-09-095-407-2
; Sequence 2, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; APPLICANT: Yip, Tai-Tung
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; TITLE OF INVENTION: Detection of Analytes
; Detection of Analytes
```

```

; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-095-407-2

Query Match          19.2%; Score 185; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      34 GHHPGHHPGHHPGHHPGHHPGHHPH 58
        |||||||
Db      1 GHHPGHHPGHHPGHHPGHHPGHHPH 25

RESULT 5
US-09-742-494-2
; Sequence 2, Application US/09742494
; Patent No. 6528320
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-2

Query Match          19.2%; Score 185; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 34 GNNPCHGNNRPHGNNRPH 58
 |
Dd 1 GNNPCHGNNRPHGNNRPH 25

```

RESULT 6
US-09-809-657-2
; Sequence 2, Application US/09809657
; Patent No. 6734022
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USF/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-2

```

Query Match	19.2%;	Score 185;	DB 4;	Length 26;
Best Local Similarity	100.0%;	Pred. NO. 9.7e-12;		
Matches 25; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	34	GHPHGHHPHGHPHGHPHGHPH	58
D_b	1	GHPHGHHPHGHPHGHPHGHPH	25

```

RESULT 7
US-09-507-323B-13
; Sequence 13, Application US/09507323B
; Patent No. 6274345
; GENERAL INFORMATION:
; APPLICANT: Choi, Jong Hyun
; APPLICANT: Lee, Sang Yup
; APPLICANT: Xu, Zhaochui
; TITLE OF INVENTION: EXPRESSION VECTORS COMPRISING A GENE
; TITLE OF INVENTION: ENCODING OMPC DERIVED FROM E. COLI AS
; TITLE OF INVENTION: MOTIF CELL SURFACE ANCHORING
; FILE REFERENCE: HYLEE39.001AUS
; CURRENT APPLICATION NUMBER: US/09/507,323B
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: KR 10-1999-0005773
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Histidine linker.
US-09-507-323B-13

```

Query Match	19.1%;	Score 184.5;	DB 3;	Length 167;
Best Local Similarity	32.7%;	Pred. No. 7.9e-11;		
Matches	48; Conservative	2; Mismatches	50; Indels	47; Gaps

[illegible]

Db 148 DHE-----HSHEH 155

RESULT 10

US-09-248-796A-24231

; Sequence 24231, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 24231

; LENGTH: 125

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-24231

Query Match 16.8%; Score 162.5; DB 4; Length 125;

Best Local Similarity 32.4%; Pred. No. 8.4e-09;

Matches 34; Conservative 2; Mismatches 22; Indels 47; Gaps 5;

QY 3 HKGHSHEQ--HPPGHHPAHHPHE-----HDTH 28

:||| | | ||| | | ||| | |

Db 20 NKGHQHNQPHPHHHHRQHHDHMKHNIYSNYPGHTYAATAPPGSPPEYGFNHSPPF 79

||| | | ||| | | ||| | |

QY 29 RQHPHGHHPHG-----HHPHGHH-----PHGHHPHGHH-PHCH 60

||| | | ||| | | ||| | |

Db 80 NAHHHHHNNRRHSFGPHHEHGHDGPMGSHGHHHHEHGGPHHH 124

RESULT 11

US-09-252-991A-28819

; Sequence 28819, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28819

; LENGTH: 225

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28819

Query Match 16.7%; Score 161.5; DB 4; Length 225;

Best Local Similarity 32.3%; Pred. No. 2e-08;

Matches 52; Conservative 3; Mismatches 49; Indels 57; Gaps 11;

QY 10 QHPHGHHPHA-----HHPHEHDTHRQHPHGHHPH-----GHHP-----HGHH---- 46

:||| | | ||| | | ||| | |

Db 38 RHPAGHRPHRLGQDHPVVRQPDHAQRPHAEPHRRGSDRIPPGGHRPDPGQRGRHDLRP 97

||| | | ||| | | ||| | |

QY 47 -PHGHHPHG-----HHPHCHDFQDYGP--CDPPP-----HNQGHCC-----HGHGP 84

||| | | ||| | | ||| | |

Db 98 RPAGDPPAGPGRGDGRDRPRPGDRDRRGAGFADRPFGALDPAHQQRHRCDDHPPGGHGHR 157

QY 85 PPG---HLRRRG--PGKGPRPFCHCRQIGSVYRLPPLRKGEV 120

| | | | | | | | | | | | | | | | | |

Db 158 VPAVLHAWRAGPAPGAGALP-----GVQGALPRRRGRV 191

| | | | | | | | | | | | | | | | | |

RESULT 12

US-09-252-991A-25291

; Sequence 25291, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25291

; LENGTH: 218

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25291

Query Match 16.4%; Score 158.5; DB 4; Length 218;

Best Local Similarity 30.7%; Pred. No. 3.8e-08;

Matches 46; Conservative 5; Mismatches 50; Indels 49; Gaps 9;

QY 1 HPHKXHSHEQHP-----HGHHPHAHHPHEHDTHRQHP---HGHHP-- 37

| | | | | | | | | | | | | | | | | |

Db 67 HLHVHALHRHHPAVHGVALHGALHHLHRS SHAHHRHA-HPHAHRRHRRHAGLGHAAHRQV 125

| | | | | | | | | | | | | | | | | |

QY 38 -----HGH---HPHGH-----HPHGHHPH--GHHPHCHDFQDYGCDPPPHNQGH 78

||| | | ||| | | ||| | |

Db 126 AALHYHGHAAVHHRRHHLGAVAEALAHVHAHRRHHLAQHAVVHRFALHG-----RHHHRHA 180

||| | | ||| | | ||| | |

QY 79 CHGHGPPPG-HLRRRGPGKGRPFHCRQIG 107

||| | | ||| | | ||| | |

Db 181 HHGHRHGGIHVHRHLHGIHHGGNSCRRSG 210

RESULT 13

US-09-949-016-8369

; Sequence 8369, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for windows Version 4.0

; SEQ ID NO 8369

; LENGTH: 955

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-8369

Query Match 16.4%; Score 158; DB 4; Length 955;

Best Local Similarity 32.0%; Pred. No. 2e-07;

Matches 49; Conservative 1; Mismatches 29; Indels 74; Gaps 12;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:51:11 ; Search time 124.225 Seconds
(without alignments)
456.700 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPHKHSHEQHPHGHPHAH.....PSFPLPHKHPLKPDNQPPF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	966	100.0	148	14 US-10-074-225A-5	Sequence 5, Appli
2	966	100.0	525	14 US-10-074-225A-1	Sequence 1, Appli
3	966	100.0	525	17 US-10-868-577A-42	Sequence 42, Appli
4	949	98.2	525	10 US-09-919-039-62	Sequence 62, Appli
5	543	56.2	526	14 US-10-074-225A-3	Sequence 3, Appli
6	325	33.6	67	15 US-10-424-599-166051	Sequence 166051,
7	302	31.3	75	9 US-09-730-379B-1	Sequence 1, Appli
8	280.5	29.0	101	14 US-10-074-225A-6	Sequence 6, Appli
9	242.5	25.1	378	14 US-10-029-386-33892	Sequence 33892, A
10	221	22.9	292	9 US-09-864-761-37944	Sequence 37944, A
11	213	22.0	38	9 US-09-730-379B-9	Sequence 9, Appli

12	213	22.0	38	9 US-09-730-379B-13	Sequence 13, Appli
13	210	21.7	82	9 US-09-864-761-33313	Sequence 33313, A
14	199.5	20.7	130	15 US-10-389-566-2423	Sequence 2423, Ap
15	199.5	20.7	133	15 US-10-389-566-2424	Sequence 2424, Ap
16	199.5	20.7	265	15 US-10-389-566-2422	Sequence 2422, Ap
17	192.5	19.9	342	15 US-10-424-599-281824	Sequence 281824,
18	191.5	19.8	156	9 US-09-864-761-41679	Sequence 41679, A
19	185	19.2	26	9 US-09-809-657-2	Sequence 2, Appli
20	185	19.2	26	9 US-09-848-512A-2	Sequence 2, Appli
21	185	19.2	26	9 US-09-742-494-2	Sequence 2, Appli
22	185	19.2	26	14 US-10-153-312A-4	Sequence 4, Appli
23	185	19.2	78	9 US-09-864-761-37352	Sequence 37352, A
24	183	18.9	96	16 US-10-425-115-191894	Sequence 191894,
25	183	18.9	213	17 US-10-871-841-2	Sequence 2, Appli
26	183	18.9	510	16 US-10-425-115-185164	Sequence 185164,
27	182	18.8	253	15 US-10-282-122A-45941	Sequence 45941, A
28	179.5	18.6	87	9 US-09-864-761-33727	Sequence 33727, A
29	179.5	18.6	87	9 US-09-864-761-34744	Sequence 34744, A
30	178.5	18.5	140	15 US-10-389-566-825	Sequence 825, App
31	176	18.2	636	16 US-10-425-115-314599	Sequence 314599,
32	175	18.1	430	15 US-10-425-114-69999	Sequence 69999, A
33	172	17.8	153	16 US-10-425-115-207055	Sequence 207055,
34	172	17.8	314	15 US-10-424-599-186228	Sequence 186228,
35	171.5	17.8	144	15 US-10-424-599-160290	Sequence 160290,
36	167.5	17.3	109	16 US-10-425-115-307018	Sequence 307018,
37	166	17.2	153	14 US-10-083-357-1014	Sequence 1014, Ap
38	164.5	17.0	695	10 US-09-890-688-130	Sequence 130, App
39	164	17.0	49	9 US-09-864-761-37882	Sequence 37882, A
40	164	17.0	108	14 US-10-029-386-31185	Sequence 31185, A
41	164	17.0	574	14 US-10-156-761-14106	Sequence 14106, A
42	162	16.8	476	15 US-10-360-849A-33	Sequence 33, Appl
43	161	16.7	371	16 US-10-425-115-320892	Sequence 320892,
44	160.5	16.6	366	15 US-10-406-686A-76	Sequence 76, Appl
45	158	16.4	139	16 US-10-437-963-122803	Sequence 122803,

ALIGNMENTS

RESULT 1
US-10-074-225A-5
; Sequence 5, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-5

Query Match	100.0%;	Score 966;	DB 14;	Length 148;
Best Local Similarity	100.0%;	Pred. No. 1.1e-70;		
Matches 148;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	HPHKHSHEQHPHGHPHAHPHEHDTHQHPHGHPHGHPHGHPHGHPHCH	60	
Db	1	HPHKHSHEQHPHGHPHAHPHEHDTHQHPHGHPHGHPHGHPHGHPHCH	60	
Qy	61	DFQDYGPCDPPPHNQHCCHGPPPGHLRRRGPCKGPRPFHCRQIGSVYRLPPLRKGEV	120	

Db 61 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120

QY 121 LPLPEANFSPFPLPHHKHPLKPDNQPPF 148

Db 121 LPLPEANFSPFPLPHHKHPLKPDNQPPF 148

RESULT 2

US-10-074-225A-1

; Sequence 1, Application US/10074225A

; Publication No. US20030082740A1

; GENERAL INFORMATION:

; APPLICANT: DONATE, Fernando

; APPLICANT: PLUNKETT, Marian L

; APPLICANT: HARRIS, Scott

; APPLICANT: MAZAR, Andrew P

; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC

; TITLE OF INVENTION: ANTI-TUMOR AGENT

; FILE REFERENCE: 38342-178463

; CURRENT APPLICATION NUMBER: US/10/074,225A

; CURRENT FILING DATE: 2002-02-14

; PRIOR APPLICATION NUMBER: US 60/268,370

; PRIOR FILING DATE: 2001-02-14

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 525

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-074-225A-1

Query Match 100.0%; Score 966; DB 14; Length 525;

Best Local Similarity 100.0%; Pred. No. 3.4e-70;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPHGHHPHAHPHEHDTHRQHPHGHHPPHGHHPHGHHPHGHHPHGHHPHCH 60

Db 350 HPHKHSHEQHPHGHHPHAHPHEHDTHRQHPHGHHPPHGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120

Db 410 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFSPFPLPHHKHPLKPDNQPPF 148

Db 470 LPLPEANFSPFPLPHHKHPLKPDNQPPF 497

RESULT 3

US-10-868-577A-42

; Sequence 42, Application US/10868577A

; Publication No. US20050032697A1

; GENERAL INFORMATION:

; APPLICANT: Alitalo et al.

; TITLE OF INVENTION: HEPARIN BINDING VEGFR-3 LIGANDS

; FILE REFERENCE: 28967/39359A

; CURRENT APPLICATION NUMBER: US/10/868,577A

; CURRENT FILING DATE: 2004-06-14

; PRIOR APPLICATION NUMBER: US 60/478,390

; PRIOR FILING DATE: 2003-06-12

; PRIOR APPLICATION NUMBER: US 10/669,176

; PRIOR FILING DATE: 2003-09-23

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 42

; LENGTH: 525

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-868-577A-42

Query Match 100.0%; Score 966; DB 17; Length 525;

Best Local Similarity 100.0%; Pred. No. 3.4e-70;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPHGHHPHAHPHEHDTHRQHPHGHHPPHGHHPHGHHPHGHHPHGHHPHCH 60

Db 350 HPHKHSHEQHPHGHHPHAHPHEHDTHRQHPHGHHPPHGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120

Db 410 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFSPFPLPHHKHPLKPDNQPPF 148

Db 470 LPLPEANFSPFPLPHHKHPLKPDNQPPF 497

RESULT 4

US-09-919-039-62

; Sequence 62, Application US/09919039

; Publication No. US20030108871A1

; GENERAL INFORMATION:

; APPLICANT: Kaser, Matthew R.

; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

; FILE REFERENCE: PA-0035 US

; CURRENT APPLICATION NUMBER: US/09/919,039

; CURRENT FILING DATE: 2002-09-09

; PRIOR APPLICATION NUMBER: 60/222,113

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 401

; SOFTWARE: PERL Program

; SEQ ID NO 62

; LENGTH: 525

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030108871A1 085596CD1

US-09-919-039-62

Query Match 98.2%; Score 949; DB 10; Length 525;

Best Local Similarity 98.6%; Pred. No. 8.1e-69;

Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPHGHHPHAHPHEHDTHRQHPHGHHPPHGHHPHGHHPHGHHPHGHHPHCH 60

Db 350 HPHKHSHEQHPHGHHPHAHPHEHDTHRQHPHGHHPPHGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 120

Db 410 DFQDYGPCDPPPHNQGHCHGPPPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFSPFPLPHHKHPLKPDNQPPF 148

Db 470 LPLPEANFSPFPLPHHKHPLKPDNQPPF 497

RESULT 5

US-10-074-225A-3

; Sequence 3, Application US/10074225A

; Publication No. US20030082740A1

; GENERAL INFORMATION:

; APPLICANT: DONATE, Fernando

; APPLICANT: PLUNKETT, Marian L

; APPLICANT: HARRIS, Scott

; APPLICANT: MAZAR, Andrew P

; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC

; TITLE OF INVENTION: ANTI-TUMOR AGENT

; FILE REFERENCE: 38342-178463

; CURRENT APPLICATION NUMBER: US/10/074,225A

; CURRENT FILING DATE: 2002-02-14

; PRIOR APPLICATION NUMBER: US 60/268,370

; PRIOR FILING DATE: 2001-02-14

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3


```

; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match      56.2%; Score 543; DB 14; Length 526;
Best Local Similarity 57.6%; Pred. No. 5.7e-36;
Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

Qy 2 PHKHSHSEQHPHGHHPHA-----HH-----PHEHDTHQHPHGHHPHGHHHP 42
Db 330 PHGHHPHGPPPHGHHPHGPPPHGHPHPPHPPHGGPPPHGHPHPPHGGPPPHGHPHPP 389

Qy 43 HGHHPHGHHPHGHCHDFQDYGPCDPPPHNQ-----HCCHGHGPPPGHLRRRGPGKGP 98
Db 390 HGHPPHGGPPPHGHPHGHFDHGCDPPSHKEGPQDLH-QHAMGPPPKHPGKRGPKGH 448

Qy 99 RPFHCRQIGSVYRLPPLRKGEVLPPLPEANFSPFLPHHKHPLKPDNQPP 148
Db 449 FPFHRRIGSVYQLPPLQKGEVLPPLPEANFPQLLLRNHTHPLKPEIQPP 498

RESULT 6
US-10-424-599-166051
; Sequence 166051, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166051
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(67)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120960C.1.pap
US-10-424-599-166051

Query Match      33.6%; Score 325; DB 15; Length 67;
Best Local Similarity 92.2%; Pred. No. 3.7e-19;
Matches 47; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HPHKHSHSEQHPHGHHPHAEHPHEHDTHQHPHGHHPHGHHPHGHHPHGHH 51
Db 17 HPHKHSHSEQHPHGHHPHAEHPHEHDTHQHPHGHHPXGNHHPHGHTPHGQH 67

RESULT 7
US-09-730-379B-1
; Sequence 1, Application US/09730379B
; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; TITLE OF INVENTION: GLYCOPROTEIN AND METHODS OF USE
; FILE REFERENCE: 955-7P/CON
; CURRENT APPLICATION NUMBER: US/09/730,379B
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match      56.2%; Score 543; DB 14; Length 526;
Best Local Similarity 57.6%; Pred. No. 5.7e-36;
Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

Qy 2 PHKHSHSEQHPHGHHPHA-----HH-----PHEHDTHQHPHGHHPHGHHHP 42
Db 330 PHGHHPHGPPPHGHHPHGPPPHGHPHPPHPPHGGPPPHGHPHPPHGGPPPHGHPHPP 389

Qy 43 HGHHPHGHHPHGHCHDFQDYGPCDPPPHNQ-----HCCHGHGPPPGHLRRRGPGKGP 98
Db 390 HGHPPHGGPPPHGHPHGHFDHGCDPPSHKEGPQDLH-QHAMGPPPKHPGKRGPKGH 448

Qy 99 RPFHCRQIGSVYRLPPLRKGEVLPPLPEANFSPFLPHHKHPLKPDNQPP 148
Db 449 FPFHRRIGSVYQLPPLQKGEVLPPLPEANFPQLLLRNHTHPLKPEIQPP 498

RESULT 6
US-10-424-599-166051
; Sequence 166051, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166051
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(67)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120960C.1.pap
US-10-424-599-166051

Query Match      33.6%; Score 325; DB 15; Length 67;
Best Local Similarity 92.2%; Pred. No. 3.7e-19;
Matches 47; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HPHKHSHSEQHPHGHHPHAEHPHEHDTHQHPHGHHPHGHHPHGHHPHGHH 51
Db 17 HPHKHSHSEQHPHGHHPHAEHPHEHDTHQHPHGHHPXGNHHPHGHTPHGQH 67

RESULT 7
US-09-730-379B-1
; Sequence 1, Application US/09730379B
; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; TITLE OF INVENTION: GLYCOPROTEIN AND METHODS OF USE
; FILE REFERENCE: 955-7P/CON
; CURRENT APPLICATION NUMBER: US/09/730,379B
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION: PUBLICATION INFORMATION
; JOURNAL: Biochemistry
; VOLUME: 25
; ISSUE: 8
; PAGES: 2220-2225
; DATE: 1986
; DATABASE ACCESSION NUMBER: GenBank/P04196
US-09-730-379B-1

Query Match      31.3%; Score 302; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 3e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 GPRPFHCRQIGSVYRLPPLRKGEVLPPLPEANFSPFLPHHKHPLKPDNQPP 148
Db 1 GPRPFHCRQIGSVYRLPPLRKGEVLPPLPEANFSPFLPHHKHPLKPDNQPP 52

RESULT 8
US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match      29.0%; Score 280.5; DB 14; Length 101;
Best Local Similarity 51.6%; Pred. No. 2.2e-15;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;

Qy 2 PHKHSHSEQHPHGHHPHA-----HH-----PHEHDTHQHPHGHHPHGHHHP 42
Db 10 PHGHHPHGPPPHGHHPHGPPPHGHPHPPHPPHGGPPPHGHPHPPHGGPPPHGHPHPP 69

Qy 43 HGHHPHGHHPHGHHPHCHDFQDYGPCDPPPH 73
Db 70 HGHPPHGGPPPHGHHPHGHPHGHGFDHGCDPPSH 100

RESULT 9
US-10-029-386-33892
; Sequence 33892, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
```

```
; SEQ ID NO 33892
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP001751.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
US-10-029-386-33892

Query Match      25.1%; Score 242.5; DB 14; Length 378;
Best Local Similarity 47.1%; Pred. No. 8.5e-12;
Matches 41; Conservative 1; Mismatches 34; Indels 11; Gaps 2;

QY 1 HPHKHSHEQHPHGHHPHAHHPHEHDTTHRQHPHGHHHPHGHH-----PHGHHPHGHHPHGH 55
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 HHHHHHHHHHHHQQHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 98

QY 56 HPHCHDFQDYGPCDPPPHNQGHCHGH 82
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 HHHHHHHHY-----HHHQQHHHHHH 119

RESULT 10
US-09-864-761-37944
; Sequence 37944, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37944
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005414.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 11
US-09-864-761-37944

Query Match      22.9%; Score 221; DB 9; Length 292;
Best Local Similarity 44.1%; Pred. No. 3.7e-10;
Matches 41; Conservative 3; Mismatches 31; Indels 18; Gaps 4;

QY 3 HKHHSHEQHPHGHHPHAHHPHEHDTTHRQHPHGHHHPHGHH-HPHGHHPHGHHPHCHD 61
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 HHHHHHDYDH-HHHHHHHYHHHHHHYHHHPHHHHHHHHHHHHHHHHHHHHHHHHHHHH- 83

QY 62 FQDYGPCDPPPHNQGHCHGHGPPPGHLRRGP 94
   ||:|||||:|||||:|||||:|||||:
Db 84 -----PHHHHHHHHH-----HRRHY 101

RESULT 11
US-09-730-379B-9
; Sequence 9, Application US/09730379B
; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; TITLE OF INVENTION: GLYCOPROTEIN AND METHODS OF USE
; FILE REFERENCE: 955-7P/CON
; CURRENT APPLICATION NUMBER: US/09/730,379B
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION: PUBLICATION INFORMATION
; JOURNAL: Biochemistry
; VOLUME: 25
; ISSUE: 8
; PAGES: 2220-2225
; DATE: 1986
; DATABASE ACCESSION NUMBER: GenBank/P04196
US-09-730-379B-9

Query Match      22.0%; Score 213; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 GPRPFHCRQIGSVYRLPPLRKGEVLPLPEANFSPFLP 134
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 GPRPFHCRQIGSVYRLPPLRKGEVLPLPEANFSPFLP 38

RESULT 12
US-09-730-379B-13
; Sequence 13, Application US/09730379B
```

```

; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; TITLE OF INVENTION: GLYCOPROTEIN AND METHODS OF USE
; FILE REFERENCE: 955-7P/CON
; CURRENT APPLICATION NUMBER: US/09/730,379B
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION: PUBLICATION INFORMATION
; JOURNAL: Biochemistry
; VOLUME: 25
; ISSUE: 8
; PAGES: 2220-2225
; DATE: 1986
; DATABASE ACCESSION NUMBER: GenBank/P04196
; US-09-730-379B-13

```

Query Match 22.0%; Score 213; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 GPRPFCHRCQIGSVYRLPPLRKGEVLPLPEANFPSFPLP 134
|||
Db 1 GPRPFCHRCQIGSVYRLPPLRKGEVLPLPEANFPSFPLP 38
|||

RESULT 13
US-09-864-761-33313
; Sequence 33313, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33313
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007263.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.1
; OTHER INFORMATION: SWISSPROT HIT: Q12329, EVALUATE 4.90e+00
US-09-864-761-33313

```

Query Match 21.7%; Score 210; DB 9; Length 82;
Best Local Similarity 56.9%; Pred. No. 9e-10;
Matches 33; Conservative 1; Mismatches 22; Indels 2; Gaps 2;

Qy	1	27
Qy	1 HPRKSHSEQHPHGHHPRHAHPHREHDTHTHQHPHGHHPRHGHHPRHGHHPRH 58	
Db		27 HMGHHHHHHYHHHDDHHHHHHHNOHQO-HHQOHHHDDH-HHHHHHHHHHHHHHHYHHRH 82

RESULT 14
US-10-389-566-2423
; Sequence 2423, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2423
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-389-566-2423

Query Match 20.7%; Score 199.5; DB 15; Length 130;
Best Local Similarity 44.6%; Pred. No. 9.7e-09;
Matches 45; Conservative 2; Mismatches 35; Indels 19; Gaps 8;

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:37:20 ; Search time 28.5301 Seconds
(without alignments)
499.124 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPHKHSHEQHPGHHPAH.....PSFPLPHHKHPLKPDNQPPF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	966	100.0	525	1	KGHUGH	histidine-rich gly
2	503.5	52.1	445	2	A60488	histidine-rich gly
3	241.5	25.0	351	1	KGZQHL	histidine-rich gly
4	217	22.5	140	2	A54523	histidine-rich pro
5	193	20.0	735	2	T45059	hypothetical prote
6	176	18.2	1571	2	T14155	zinc finger protei
7	173.5	18.0	471	2	T33997	hypothetical prote
8	171.5	17.8	657	2	A29454	knob-associated hi
9	170.5	17.7	380	2	H83994	hypothetical prote
10	167	17.3	361	2	F87286	cation efflux fami
11	166.5	17.2	369	2	I78877	cAMP responsive el
12	166.5	17.2	508	2	A45477	cAMP response elem
13	166.5	17.2	515	2	T23089	hypothetical prote
14	166.5	17.2	1378	2	T30173	zinc finger protei
15	162	16.8	436	2	I49714	MHC H-2K/t-w5-link
16	160	16.6	427	2	A32372	female-specific do
17	160	16.6	549	2	B32372	male-specific doub
18	158	16.4	428	2	T27544	zinc resistance pr
19	155.5	16.1	774	2	T39539	alpha-amylase homo
20	155	16.0	199	2	T48099	hypothetical prote
21	149.5	15.5	659	2	A36664	S59/2 homeotic pr
22	149.5	15.5	1002	2	T43236	carboxypeptidase C
23	148.5	15.4	242	2	H82061	hypothetical prote
24	147	15.2	373	2	AD0262	probable membrane
25	147	15.2	375	2	AB3110	conserved hypothet
26	147	15.2	375	2	B98177	hypothetical prote
27	145.5	15.1	654	2	B71623	knob-associated Hi
28	145.5	15.1	1398	2	T13741	hypothetical prote
29	144.5	15.0	60	2	C64698	probable histidine

30	143.5	14.9	852	2	A34373	histidine-rich cal
31	143	14.8	731	1	JC2464	probable copper-tr
32	143	14.8	940	2	JE0291	FB19 protein - hum
33	142.5	14.8	102	2	T30119	hypothetical prote
34	142.5	14.8	270	2	A26480	knob protein - mal
35	142.5	14.8	473	2	A54494	knob-associated hi
36	142.5	14.8	634	2	A28412	histidine-rich pro
37	142.5	14.8	634	2	A54495	knob protein precu
38	142	14.7	301	2	S31782	HRPII protein - ma
39	142	14.7	997	2	T28872	hypothetical prote
40	141	14.6	905	2	AC2680	hypothetical prote
41	141	14.6	916	2	A97462	hypothetical prote
42	140	14.5	297	2	S23737	proline-rich prote
43	140	14.5	309	2	A25942	histidine/alanine-
44	139	14.4	130	2	S14983	extensin class I (
45	139	14.4	330	2	S22140	nodulin Enod2 - Se

ALIGNMENTS

RESULT 1

KGHUGH

histidine-rich glycoprotein precursor - human

N;Alternate names: HRG

C;Species: Homo sapiens (man)

C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004

C;Accession: A01287; S29669

R;Koide, T.; Foster, D.; Yoshitake, S.; Davie, E.W.

Biochemistry 25, 2220-2225, 1986

A;Title: Amino acid sequence of human histidine-rich glycoprotein derived from the nucle

A;Reference number: A01287; MUID:86216149; PMID:3011081

A;Accession: A01287

A;Molecule type: mRNA

A;Residues: 1-525 <KOI>

A;Cross-references: UNIPROT:P04196; GB:AB005803; NID:G2280513; PIDN:BAA21613.1; PID:G228

R;Hennis, B.; Havelaar, A.; Kluft, C.

submitted to the EMBL Data Library, October 1991

A;Description: PCR detection of a dinucleotide repeat in the human histidine-rich glycop

A;Reference number: S29669

A;Accession: S29669

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 214-247 <HEN>

A;Cross-references: EMBL:Z17218; NID:G32453; PIDN:CAA78925.1; PID:G32454

C;Comment: Although its physiological function is not yet known, HRG does bind heme, dyes
din, and the lysine-binding site of plasminogen. On the basis of its homology with HMW k
lood coagulation cascade.

C;Comment: The amino half of this protein is homologous to the first two cystatin-like d
ould not have inhibitory activity.

C;Comment: In addition to having a high histidine and proline content, this protein has n
e-rich' region.

C;Genetics:

A;Gene: GDB:HRG

A;Cross-references: GDB:120055; OMIM:142640

A;Map position: 3q27-3q27

C;Superfamily: histidine-rich glycoprotein; cystatin homology

C;Keywords: duplication; glycoprotein; heparin binding; tandem repeat

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-525/Product: histidine-rich glycoprotein #status predicted <MAT>

F;19-131/Domain: cystatin homology <CY1>

F;140-246/Domain: cystatin homology <CY2>

F;276-321/Region: proline-rich

F;348-437/Region: histidine-rich

F;351-497/Region: proline-rich

F;63,125,344,345/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;78-89,105-126,218-241/Disulfide bonds: #status predicted

Query Match 100.0%; Score 966; DB 1; Length 525;

Best Local Similarity 100.0%; Pred. No. 1.4e-74;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPGHHPAHPEHDTHRQHPHGHHPHGHHPHGHHPHCH 60

Db 350 HPHKHSHEQHPGHGHPHAHPHEHDTHQHPGHHPHGHHHPGHHPGHHPHCH 409

QY 61 DFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGPGKGPRPFHCRQIGSVYRLPPLRKGEV 120

Db 410 DFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGPGKGPRPFHCRQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFPSPFLPHHKHPLKPDNQPPF 148

Db 470 LPLPEANFPSPFLPHHKHPLKPDNQPPF 497

RESULT 2

A60488

histidine-rich glycoprotein - bovine (fragments)

N;Alternate names: autorosette inhibition factor

C;Species: Bos primigenius taurus (cattle)

C;Date: 19-Mar-1993 #sequence revision 23-Mar-1995 #text_change 07-Jul-1995

C;Accession: S35687; JC2196; A60488

R;Sorensen, C.B.; Krogh-Pedersen, H.; Petersen, T.E.

FEBS Lett. 328, 285-290, 1993

A;Title: Determination of the disulphide bridge arrangement of bovine histidine-rich gly

A;Reference number: S35687; MUID:93351678; PMID:8348977

A;Accession: S35687

A;Molecule type: protein

A;Residues: 1-25;26-52;57-75;82-88;95-119;146-173;175-206;210-309;313-445 <SOR>

A;Note: 355-Gln and 368-Tyr were also found

R;Halkier, T.; Andersen, H.; Vestergaard, A.; Magnusson, S.

Biochem. Biophys. Res. Commun. 200, 78-82, 1994

A;Title: Bovine histidine-rich glycoprotein is a substrate for bovine plasma factor XIII

A;Reference number: JC2196; MUID:94220160; PMID:7909439

A;Accession: JC2196

A;Molecule type: protein

A;Residues: 1-23;35-54,'VK',57-101,'R','TVGEYTEG',116,'N',118,'R',120-136;137-145;150-20

A;Experimental source: plasma

R;Vestergaard, A.B.; Andersen, H.F.; Magnusson, S.; Halkier, T.

Thromb. Res. 60, 385-396, 1990

A;Title: Histidine-rich glycoprotein inhibits contact activation of blood coagulation.

A;Reference number: A60488; MUID:91196010; PMID:2084959

A;Accession: A60488

A;Molecule type: protein

A;Residues: 1-6,'X',8-15 <VES>

C;Comment: This protein is a single-chained plasma protein which participates in transgl

C;Superfamily: histidine-rich glycoprotein; cystatin homology

C;Keywords: glycoprotein; plasma

F;2-113/Domain: cystatin homology <CY1>

F;122-207/Domain: cystatin homology (fragments) <CY2>

F;7-424,60-71,87-108,165-346,180-203,258-288/Disulfide bonds: #status experimental

F;74,107,164,266/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 52.1%; Score 503.5; DB 2; Length 445;

Best Local Similarity 49.2%; Pred. No. 1.7e-35;

Matches 91; Conservative 15; Mismatches 26; Indels 53; Gaps 6;

QY 1 HPHKHS-----HEQH-----PH-----GH-----PHAHHPH 23

Db 249 HPHESYNFRCPPLLEHKNHSDSPPFQARAPLPFPPGLRCPHPPFGTKGNRRPPHDSSD 308

QY 24 EHDTHRQHPGHHPGHHPGHHPGHHPGHHPCHDFQYGPCDPPPHNQHCCHG 83

Db 309 EH-----HNHPHGHPGHHPGHHPGHHPDNDFYDHGCPDPPPHR----- 352

QY 84 PPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEVLPPEANFPSPFLPHHKHPLKPD 143

Db 353 PPRHSKERGPGKGRFHFWRPTGYIHRPLSLKGEVLPPEANFPSPFLPNHNNPLQPE 412

QY 144 NQPEP 148

Db 413 IQAFP 417

RESULT 3

KGZQHL

histidine-rich glycoprotein precursor - Plasmodium lophurae

C;Species: Plasmodium lophurae

C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004

C;Accession: A22692

R;Ravetch, J.V.; Feder, R.; Pavlovic, A.; Blobel, G.

Nature 312, 616-620, 1984

A;Title: Primary structure and genomic organization of the histidine-rich protein of the

A;Reference number: A22692; MUID:85061618; PMID:6095114

A;Accession: A22692

A;Molecule type: DNA

A;Residues: 1-351 <RAV>

A;Cross-references: UNIPROT:P04929; GB:X01469; NID:g9997; PIDN:CAA25698.1; PID:g9999

C;Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats,

C;Genetics:

A;Introns: 23/3

C;Superfamily: plasmodium histidine-rich protein

C;Keywords: glycoprotein; tandem repeat

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-47/Domain: propeptide #status predicted <PRO>

F;48-351/Product: histidine-rich glycoprotein #status predicted <MAT>

F;59-74,75-90/Region: 16-residue repeats

F;91-107,108-123/Region: 17-residue repeats

F;124-138,139-153/Region: 15-residue repeats

F;173-301,312-331/Region: 10-residue repeats

F;40/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 25.0%; Score 241.5; DB 1; Length 351;

Best Local Similarity 47.7%; Pred. No. 2e-13;

Matches 42; Conservative 1; Mismatches 34; Indels 11; Gaps 3;

QY 2 PHKHSHEQHPHGHH-----PHAHPHEH--DTHRQHPGHHPGHHPGHHPGHHPHG 54

Db 184 PHHHHHHHAPHHHHHHHHHAPHHHHHHHHHAPHHHHHHHHHGHGHHHHHHHHHG 243

QY 55 HHPHCHDFQYGPCDPPPHNQHCCHG 82

Db 244 HHHHHHHHD-----AHHHHHHHDAHHH 267

RESULT 4

A54523

histidine-rich protein - Plasmodium lophurae (fragment)

C;Species: Plasmodium lophurae

C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 09-Jul-2004

C;Accession: A54523

R;Irving, D.O.; Cross, G.A.M.; Feder, R.; Wallach, M.

Mol. Biochem. Parasitol. 18, 223-234, 1986

A;Title: Structure and organization of the histidine-rich protein gene of Plasmodium lopl

A;Reference number: A54523; MUID:86174893; PMID:3007981

A;Accession: A54523

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-140 <IRV>

A;Cross-references: UNIPROT:Q26056; GB:M15317; NID:g160331; PIDN:AAA29616.1; PID:g552196

C;Superfamily: plasmodium histidine-rich protein

C;Keywords: tandem repeat

Query Match 22.5%; Score 217; DB 2; Length 140;

Best Local Similarity 39.6%; Pred. No. 1e-11;

Matches 38; Conservative 2; Mismatches 34; Indels 22; Gaps 2;

QY 2 PHKHSHEQHPHGHPHAHPHEHDTHQHPGHHPGHHPGHG-----H 46

Db 8 PHHHHHHHAPHHHHHHHHHAPHHHHHHHHHAPHHHHHHHHHAPHHHHHHHPWFH 67

QY 47 PHGHPGHHPHCHDFQYGPCDPPPHNQHCCHG 82

Db 68 HHHHPWFHHHHHHDAHH-----HHHHDAHH 96

RESULT 5

T45059

hypothetical protein Y39B6B.gg [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45059
R;Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johns B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D. Nature 368, 32-38, 1994
A;Authors: Showkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Stock, L.; Wilkinson-Sproat, J.; Wohldman, P.
A;Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A;Reference number: S43531; MUID:94150718; PMID:7906398
A;Accession: T45059
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-735 <WIL>
A;Cross-references: UNIPROT:Q9NES7; EMBL:AL132896; NID:g6434440; PIDN:CAB60938.1; PID:g6
A;Experimental source: clone Y39B6B
C;Genetics:
A;Map position: 3
A;Introns: 18/1; 69/1
A;Note: Y39B6B.gg

Query Match 20.0%; Score 193; DB 2; Length 735;
Best Local Similarity 27.9%; Pred. No. 5.2e-09;
Matches 60; Conservative 4; Mismatches 65; Indels 86; Gaps 11;

QY 1 HPHKHHS-HEQHPHGHHPAHHPH-----EHDTHRQH 31
Db 434 HAPAHGHSHESGHHSPAHGHGHEHHHAPAHGHGHEHHHAPAHGHGHEGTHGHG 493

QY 32 PHGHHP---HGHH-----PHGHH-----PHGHH-----PHCHDF 62
Db 494 HGSHHSPAHGHGHEHHHAPAHGHGHEGTHGHGHEHHHAPAHGHGHEGTHGHG 553

QY 63 QDYGPC-----DPPPHNQHCCHG-HGPPPGHLRRRPG-----KGPRPFHC 103
Db 554 SHHSPAHGHGHEHHHAPAHGHGHHGHSVHHGHGHESHGHHAPAHGHGHGEGVH 613

QY 104 RQIGSVYRLPPLRKGEVLPPLPEANFSPFPLPHKH 138
Db 614 GHGAGYGAHHGHGGA-----HHHHAPHHEH 639

RESULT 6
T14155
zinc finger protein Peg3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14155
R;Kuroiwa, Y.; Kaneko-Ishino, T.; Kagitani, F.; Kohda, T.; Li, L.L.; Tada, M.; Suzuki, R Nature Genet. 12, 186-190, 1996
A;Title: Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein.
A;Reference number: Z17892; MUID:96154192; PMID:8563758
A;Accession: T14155
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-1571 <KUR>
A;Cross-references: UNIPROT:O54978; EMBL:AF038939; NID:g2791677; PID:g2791678; PIDN:AAB9
C;Genetics:
A;Gene: Peg3
A;Map position: 7
C;Keywords: zinc finger

Query Match 18.2%; Score 176; DB 2; Length 1571;
Best Local Similarity 36.5%; Pred. No. 2.9e-07;
Matches 31; Conservative 5; Mismatches 45; Indels 4; Gaps 1;

QY 2 PHKHHSHEQHPGHHPHAHHPHEHDTHRQHPGHHPHGHHHPGHHPGHHPGHCHD 61
Db 977 PLVQEMRSEEPHDDKPHGQEPHDDKPHGQEPHGDDEPHGQEPHGDPEPHDK 1036

QY 62 FQDYGPCDPPPHNQGHCHGHGPP 86

Db 1037 PIDQEMRSEEPHSE----ESHGDEP 1057

RESULT 7
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33997
R;Pauley, A.; Scheet, P.; Harper, M.
A;Submitted to the EMBL Data Library, February 1999
A;Description: The sequence of C. elegans cosmid W03G1.
A;Reference number: Z21454
A;Accession: T33997
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-471 <PAU>
A;Cross-references: UNIPROT:Q9UAY0; EMBL:AF125964; PIDN:AAD14753.1; GSPDB:GN00022; CESP:Y
A;Experimental source: strain Bristol N2; clone W03G1
C;Genetics:
A;Gene: CESP:W03G1.5
A;Map position: 4

Query Match 18.0%; Score 173.5; DB 2; Length 471;
Best Local Similarity 40.6%; Pred. No. 1.5e-07;
Matches 39; Conservative 2; Mismatches 36; Indels 19; Gaps 6;

QY 3 HKHHSHEQHPGH---HPHAHHPHEHDTHRQHPGHHPHGHHP---HGHHPHGH----- 51
Db 369 HGSRSHPRGHGGGRHGPPHPCGRHG-HHGPPHHHHHDSRSPSRHGHHHHHHGCRPP 427

QY 52 PHGHHPHCHDFQDYGPCDP-----PPHNQGHCCGH 82
Db 428 PHGHHHFFPF--WPCCPPPPFPFPPHRRGCHHH 461

RESULT 8
A29454
knob-associated histidine-rich protein precursor - malaria parasite (Plasmodium falcipar C;Species: Plasmodium falciparum
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: A29454
R;Triglia, T.; Stahl, H.D.; Crewther, P.E.; Scanlon, D.; Brown, G.V.; Anders, R.F.; Kemp, EMBO J. 6, 1413-1419, 1987
A;Title: The complete sequence of the gene for the knob-associated histidine-rich protein
A;Reference number: A29454; MUID:87275836; PMID:3301325
A;Accession: A29454
A;Molecule type: DNA
A;Residues: 1-657 <TRI>
A;Cross-references: UNIPROT:P06719; GB:Y00060; NID:g9908; PIDN:CAA68268.1; PID:g9909
C;Superfamily: knob-associated histidine-rich protein

Query Match 17.8%; Score 171.5; DB 2; Length 657;
Best Local Similarity 39.0%; Pred. No. 3.1e-07;
Matches 32; Conservative 5; Mismatches 32; Indels 13; Gaps 2;

QY 7 SHEQHPGHHPHAHHPHEHDTHRQHPGHHPHGHHPHGHHPGHHP-----HG-----HHPH 53
Db 57 AQKQHEHHHHHHHHHHQHQAHPQAHHHHHHGVEVNHQAPOVHQVQVHQDQAHHHH 116

QY 54 GHHPHCHDFQDYGPCDPPPHNQ 75
Db 117 HHHHLHPQPPQGTVANPPPSNE 138

RESULT 9
H83994
hypothetical protein BH2760 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83994
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran Nucleic Acids Res. 28, 4317-4331, 2000

RESULT 13
T23089
hypothetical protein H13N06.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23089
R;Lennard, N.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z19673
A;Accession: T23089
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-515 <WIL>
A;Cross-references: UNIPROT:O9XTQ7; EMBL:Z99942; PIDN:CABI7070.1; GSPDB:GN00028; CESP:H1
A;Experimental source: clone H13N06
C;Genetics:
A;Gene: CESP:H13N06.5
A;Map position: X
A;Introns: 118/1; 156/2; 182/2; 306/1

Query Match 17.2%; Score 166.5; DB 2; Length 515;
Best Local Similarity 28.5%; Pred. No. 6.Se-07;
Matches 39; Conservative 6; Mismatches 29; Indels 63; Gaps 7;

QY 5 HHSHEQHPHGHPAH-----HPHEHDTHRQHPHGHPHG 39
Db 92 HHQGHGAHGGHGHADADGGCPYAKAAABAAATAAHDHGHAHDHGHADHG-HAHD 150

QY 40 HHPHGH-----HPHGHHPHGHCHDFQDYGPCDPPPHNQHCCHGHPPPGHLRRRGP 94
Db 151 HHGSHDEEDHHG-HAHDHHGSH--EDHG-----HSHGAE----- 185

QY 95 GKGRPRPFHCRQIGSVYR 111
Db 186 -----SAKQVGDEYQ 195

RESULT 14
T30173
zinc finger protein Pw1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30173
R;Relaix, F.; Weng, X.; Marazzi, G.; Yang, E.; Copeland, N.; Jenkins, N.; Spence, S.E.;
Dev. Biol. 177, 383-396, 1996
A;Title: Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages.
A;Reference number: Z20754; MUID:96400442; PMID:8806818
A;Accession: T30173
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1378 <REL>
A;Cross-references: UNIPROT:Q61138; EMBL:U48804; NID:g1197856; PID:g1197857; PIDN:AAC527
A;Experimental source: strain Sv129; limb bud
C;Genetics:
A;Map position: 7

Query Match 17.2%; Score 166.5; DB 2; Length 1378;
Best Local Similarity 36.0%; Pred. No. 1.6e-06;
Matches 31; Conservative 5; Mismatches 45; Indels 5; Gaps 2;

QY 2 PHKHSHEQHPHGHPAHHPHEH-DTHRQHPHGHPHGHPHGHPHGHPHGHPHGHPHCH 60
Db 783 PLDQEMRSEEPHDDKPHGQEPHDDMRPHGQEPHDDPHGQEPHDEPHGQEPHGDPEPHDX 842

QY 61 DFQDYGPCDPPPHNQHCCHGHP 86
Db 843 EPIDQEMRSEEPHSE----ESHGDEP 864

RESULT 15
I49714

MHC H-2K/t-w5-linked ORF precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I49714
R;Han, S.
Mol. Cell. Biol. 10, 138-145, 1990
A;Title: A putative transmembrane protein with histidine-rich charge clusters encoded in
A;Reference number: I49714; MUID:90097821; PMID:2294398
A;Accession: I49714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-436 <RES>
A;Cross-references: GB:M32010; NID:g193738; PIDN:AAA37767.1; PID:g309286

Query Match 16.8%; Score 162; DB 2; Length 436;
Best Local Similarity 42.9%; Pred. No. 1.3e-06;
Matches 39; Conservative 2; Mismatches 32; Indels 18; Gaps 8;

QY 1 HPHKH-HSHEQHPHGHPHAPHHPHEHDTHRQHPHG-----HPHGH-HPHGHHPHGHPH 53
Db 43 HGHSHGSHEDFHGHSHGSH----EDFHGHGHTHESIWHGHAHSHDHGHSREELH-H 97

QY 54 GH-HPHCHDFQDYGPCDPPPHNQHCCHG 83
Db 98 GHSHGSHSDSLHHG-----GHGHAHREHSHG 123

Search completed: June 15, 2005, 13:51:55
Job time : 30.5301 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:32:29 ; Search time 133.141 Seconds
(without alignments)
569.231 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPHKHSHSEQHPHGHPHAH.....PSFPLPHKHPLKPDNQPPF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	966	100.0	525	1	HRG_HUMAN	P04196 homo sapien
2	949	98.2	312	2	Q68DR3	Q68dr3 homo sapien
3	645.5	66.8	525	2	Q99PS5	Q99ps5 mus musculu
4	645.5	66.8	525	2	Q99PS6	Q99ps6 mus musculu
5	645.5	66.8	525	2	Q9ESB3	Q9esb3 mus musculu
6	645.5	66.8	536	2	Q6YK32	Q6yk32 mus musculu
7	645.5	66.8	536	2	Q6YKA2	Q6yka2 mus musculu
8	644	66.7	525	2	Q99PS8	Q99ps8 rattus norv
9	644	66.7	546	2	Q9BGU1	Q9bgu1 bos taurus
10	626	64.8	515	2	Q99PS7	Q99ps7 rattus norv
11	614.5	63.6	510	2	Q9ESB2	Q9esb2 rattus norv
12	543	56.2	526	1	HRG_RABIT	Q28640 oryctolagus
13	489	50.6	396	1	HRG_BOVIN	P33433 bos taurus
14	281.5	29.1	233	2	Q8MP30	Q8mp30 dictyosteli
15	241.5	25.0	351	1	HRPX_PLALO	P04929 plasmodium
16	223.5	23.1	294	2	Q9KI87	Q9ki87 bacillus ce
17	217	22.5	140	2	Q26056	Q26056 plasmodium
18	216.5	22.4	296	2	Q9KI89	Q9ki89 bacillus ce
19	213.5	22.1	465	2	Q801E5	Q801e5 xenopus lae
20	208	21.5	293	2	Q632V3	Q632v3 bacillus ce
21	206	21.3	248	2	Q6HCE3	Q6hce3 bacillus th
22	205.5	21.3	273	2	Q9KI88	Q9ki88 bacillus ce
23	205.5	21.3	308	2	Q72YZ0	Q72yz0 bacillus ce
24	205.5	21.3	462	2	Q7SYH2	Q7syh2 xenopus lae
25	205.5	21.3	462	2	Q7ZY91	Q7zy91 xenopus lae
26	205	21.2	295	2	Q816T1	Q816t1 bacillus ce
27	204.5	21.2	479	2	Q8T0B3	Q8t0b3 drosophila
28	204.5	21.2	783	2	Q9V5Z7	Q9vsz7 drosophila
29	202.5	21.0	437	2	Q6NBF4	Q6nbf4 rhodospheudo
30	199.5	20.7	130	2	Q9KI77	Q9ki77 bacillus an
31	199.5	20.7	133	2	Q9KI76	Q9ki76 bacillus an

32	199.5	20.7	265	2	Q9KI91	Q9ki91 bacillus an
33	198.5	20.5	474	2	Q6DIT5	Q6dit5 xenopus tro
34	196.5	20.3	449	1	CSUP_DROME	Q9v3a4 drosophila
35	193	20.0	735	2	Q9NES7	Q9nes7 caenorhabdi
36	191.5	19.8	341	2	Q9VV44	Q9vv44 drosophila
37	191.5	19.8	356	2	Q7K1P4	Q7klp4 drosophila
38	191.5	19.8	485	2	Q76876	Q76876 drosophila
39	190.5	19.7	1131	2	Q75DU5	Q75dj5 ashbya goes
40	188	19.5	403	2	Q98H88	Q98h88 rhizobium l
41	183	18.9	213	2	Q6EI04	Q6ei04 cucurbita m
42	182	18.8	118	2	Q9KI82	Q9ki82 bacillus an
43	182	18.8	121	2	Q9KI80	Q9ki80 bacillus an
44	182	18.8	124	2	Q9KI78	Q9ki78 bacillus an
45	182	18.8	124	2	Q9KI79	Q9ki79 bacillus an

ALIGNMENTS

RESULT 1

ID	HRG_HUMAN	STANDARD;	PRT;	525 AA.
AC	P04I96;			
DT	20-MAR-1987 (Rel. 04, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG).			
DE	Name=HRG;			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=86216149; PubMed=3011081;			
RA	Koide T., Foster D.C., Yoshitake S., Davie E.W.;			
RT	"Amino acid sequence of human histidine-rich glycoprotein derived from the nucleotide sequence of its cDNA.";			
RL	Biochemistry 25:2220-2225(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;			
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 214-247 FROM N.A.			
RX	MEDLINE=94245171; PubMed=8188234;			
RA	Hennis B.C., Frants R.R., Bakker E., Vossen R.H., van der Poort E.W.,			
RA	Blonden L.A., Cox S., Khan P.M., Spurr N.K., Klufft C.;			
RT	"Evidence for the absence of intron H of the histidine-rich glycoprotein (HRG) gene: genetic mapping and in situ localization of			
RT	HRG to chromosome 3q28-q29.";			
RL	Genomics 19:195-197(1994).			
RN	[4]			
RP	SEQUENCE OF 19-27.			
RC	TISSUE=Plasma;			
RX	MEDLINE=93092937; PubMed=1459097;			
RA	Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,			
RA	Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,			
RA	Hochstrasser D.F.;			
RT	"Plasma protein map: an update by microsequencing.";			
RL	Electrophoresis 13:707-714(1992).			
CC	-I- FUNCTION: The physiological function is not yet known. It binds heme, dyes and divalent metal ions. It can inhibit rosette formation and is known to interact with heparin, thrombospondin, and the lysine-binding site of plasminogen. On the basis of its homology with HMW kininogen, the His-rich region of this protein may mediate the contact activation phase of intrinsic blood coagulation cascade.			
CC	-I- SUBCELLULAR LOCATION: Secreted.			
CC	-I- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.			
CC	-I- DOMAIN: In addition to having a high His and Pro content, this			

```
CC      protein has many internal repeats. 12 tandem repetitions of a 5-
CC      residue sequence (GHHPH, consensus) form a histidine-rich region.
CC      -!- SIMILARITY: Contains 2 cystatin-like domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M13149; AAA52694.1; -.
CC      EMBL; AB005803; BAA21613.1; -.
CC      EMBL; Z17218; CAA78925.1; -.
CC      PIR; A01287; KGHUGH.
CC      SWISS-2DPAGE; P04196; HUMAN.
CC      Genew; HGNC:5181; HRG.
CC      MIM; 142640; -.
CC      InterPro; IPR000010; Prot_inh_cystat.
CC      Pfam; PF00031; Cystatin; 1.
KW      Direct protein sequencing; Glycoprotein; Heparin-binding;
KW      Polymorphism; Repeat; Signal.
FT      SIGNAL          1      18      Histidine-rich glycoprotein.
FT      CHAIN           19      525      Cystatin-like 1.
FT      DOMAIN          19      136      Cystatin-like 2.
FT      DOMAIN          137      254      Pro-rich.
FT      DOMAIN          276      321      His/Pro-rich.
FT      DOMAIN          350      497      By similarity.
FT      DISULFID         24      504      By similarity.
FT      DISULFID         78      89      By similarity.
FT      DISULFID        105      126      By similarity.
FT      DISULFID        203      417      By similarity.
FT      DISULFID        218      241      By similarity.
FT      CARBOHYD         63      63      N-linked (GlcNAc. .) (Potential).
FT      CARBOHYD         87      87      N-linked (GlcNAc. .) (Potential).
FT      CARBOHYD        125      125      N-linked (GlcNAc. .) (Potential).
FT      CARBOHYD        344      344      N-linked (GlcNAc. .) (Potential).
FT      CARBOHYD        345      345      N-linked (GlcNAc. .) (Potential).
FT      VARIANT         204      204      P -> S (in dbSNP:3181917).
FT                                     /FTId=VAR 014528.
FT      SEQUENCE        525 AA; 59578 MW; A2B124D6CE93114F CRC64;
SQ      Query Match          100.0%; Score 966; DB 1; Length 525;
      Best Local Similarity 100.0%; Pred. No. 1.2e-67;
      Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPHKHSHEQHHPHAAHPHEHDTHQHPHGHHPHGHHPHGHHPHGHHPHCH 60
      |||||||
Db      350 HPHKHSHEQHHPHAAHPHEHDTHQHPHGHHPHGHHPHGHHPHGHHPHCH 409

QY      61 DFQDYGPCDPPPHNQGHCCGHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 120
      |||||||
Db      410 DFQDYGPCDPPPHNQGHCCGHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 469

QY      121 LPLPEANFPSPPLPHHKHPLKPDNQPPF 148
      |||||||
Db      470 LPLPEANFPSPPLPHHKHPLKPDNQPPF 497

RESULT 2
Q68DR3
ID      Q68DR3          PRELIMINARY; PRT; 312 AA.
AC      Q68DR3;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE      Hypothetical protein DKFp779H1622 (Fragment).
GN      Name=DKFp779H1622;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
```

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RG      The German cDNA Consortium;
RA      Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA      Fobo G., Han M., Wiemann S.;
RL      Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.
DR      EMBL; CR749302; CAH18157.1; -.
KW      Hypothetical protein.
FT      NON_TER         1
SQ      SEQUENCE        312 AA; 35124 MW; 08522797450AFA1F CRC64;

      Query Match          98.2%; Score 949; DB 2; Length 312;
      Best Local Similarity 98.6%; Pred. No. 1.6e-66;
      Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 HPHKHSHEQHHPHAAHPHEHDTHQHPHGHHPHGHHPHGHHPHGHHPHCH 60
      |||||||
Db      137 HPHKHSHEQHHPHAAHPHEHDTHQHPHGHHPHGHHPHGHHPHGHHPHCH 196

QY      61 DFQDYGPCDPPPHNQGHCCGHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 120
      |||||||
Db      197 DFQDYGPCDPPPHNQGHCCGHGPPPGHLLRRRGPGKGRPFHCRQIGSVYRLPLRKGEV 256

QY      121 LPLPEANFPSPPLPHHKHPLKPDNQPPF 148
      |||||||
Db      257 LPLPEANFPSPPLPHHKHPLKPDNQPPF 284

RESULT 3
Q99PS5
ID      Q99PS5          PRELIMINARY; PRT; 525 AA.
AC      Q99PS5;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE      Histidine-rich glycoprotein.
GN      Name=Hrg; Synonyms=MMHRG;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Tsuchida N., Wakabayashi S., Jahnen-Dechent W., Koide T.;
RL      Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Liver;
RX      MEDLINE=22388257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA      Jones S.J., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Liver;
```

RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055898; BAB33095.1; -.
DR EMBL; BC011168; AAH11168.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59162 MW; A83E93A439CFB126 CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 525;
Best Local Similarity 69.5%; Pred. No. 1.2e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

Qy 1 HPHKHSHEQHHPGHHPHAHHPHEHDTHRQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 60
Db 347 HPHSHHPPGGHSHGHHPGHHPHSHHSHGHHPGHHPGHHPGHHPGHHPGHHPGHHPGH 406
Qy 61 DFQDYGPCDPPPHNQ---GHCCHGHGPPPGHLLRRRGKGPFPFHCRCQIGSVYRLPPLRK 117
Db 407 DFLDYGPCDPPSNSQELKGQYHRGYGPPPHGSHSRKRGKGLFPFHHQIGYVYRLPPLNI 466
Qy 118 GEVLTPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 467 GEVLTPLPEANFPSPFLPNCNRSLOPEIQPPF 497

RESULT 4
Q99PS6 PRELIMINARY; PRT; 525 AA.
AC Q99PS6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg; Synonyms=MMHRG;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055897; BAB33094.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59090 MW; A83E93A439CFB3AC CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 525;
Best Local Similarity 69.5%; Pred. No. 1.2e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

Qy 1 HPHKHSHEQHHPGHHPHAHHPHEHDTHRQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 60
Db 347 HPHSHHPPGGHSHGHHPGHHPHSHHSHGHHPGHHPGHHPGHHPGHHPGHHPGHHPGH 406
Qy 61 DFQDYGPCDPPPHNQ---GHCCHGHGPPPGHLLRRRGKGPFPFHCRCQIGSVYRLPPLRK 117
Db 407 DFLDYGPCDPPSNSQELKGQYHRGYGPPPHGSHSRKRGKGLFPFHHQIGYVYRLPPLNI 466
Qy 118 GEVLTPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 467 GEVLTPLPEANFPSPFLPNCNRSLOPEIQPPF 497

RESULT 5

Q9ESB3 PRELIMINARY; PRT; 525 AA.
AC Q9ESB3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=20307726; PubMed=10849117;
RA Hulett M.D., Parish C.R.;
RT "Murine histidine-rich glycoprotein: cloning, characterization and
RT cellular origin."
RL Immunol. Cell Biol. 78:280-287(2000).
DR EMBL; AF194028; AAG28416.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59132 MW; 6E55F2A439CFB123 CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 525;
Best Local Similarity 69.5%; Pred. No. 1.2e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;
Qy 1 HPHKHSHEQHHPGHHPHAHHPHEHDTHRQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 60
Db 347 HPHSHHPPGGHSHGHHPGHHPHSHHSHGHHPGHHPGHHPGHHPGHHPGHHPGHHPGH 406
Qy 61 DFQDYGPCDPPPHNQ---GHCCHGHGPPPGHLLRRRGKGPFPFHCRCQIGSVYRLPPLRK 117
Db 407 DFLDYGPCDPPSNSQELKGQYHRGYGPPPHGSHSRKRGKGLFPFHHQIGYVYRLPPLNI 466
Qy 118 GEVLTPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 467 GEVLTPLPEANFPSPFLPNCNRSLOPEIQPPF 497

RESULT 6
Q6YK32 PRELIMINARY; PRT; 536 AA.
AC Q6YK32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Histidine-rich glycoprotein HRG.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RA Hsu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137504; AAN27996.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 536 AA; 60439 MW; CCF14FC08DD53D7C CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 536;
Best Local Similarity 69.5%; Pred. No. 1.3e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 1 HPHKHHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHHPHGHHHPHCH 60
Db 358 HPHSHHPGHHSHGHHPHGHHPHSHSHSGHHPPGHHHPHGHHHPHGHHHPHGHHHPHCH 417

QY 61 DFQDYGPCDPPPHNQ--GHCCHGHGPPPGHLRRRPGKGPFPFHCRCQIGSVYRLPPLRK 117
Db 418 DFLDYGPCDPPSNSQELKGQYHRGYGPPPHGHSRKRGPGLFPFHHQQIGYVYRLPPLNI 477

QY 118 GEVLPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 478 GEVLTLPEANFPSPFLPNCNRSLQPEIQPPF 508

RESULT 7
Q6YKA2
ID Q6YKA2 PRELIMINARY; PRT; 536 AA.
AC Q6YKA2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Histidine-rich glycoprotein Hrg.
GN Name=Hrg1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH/Ola;
RA Hsu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AY135662; AAN10183.1; -.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 536 AA; 60492 MW; 71CE4FC6DF3A3D72 CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 536;
Best Local Similarity 69.5%; Pred. No. 1.3e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 1 HPHKHHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHHPHGHHHPHCH 60
Db 358 HPHSHHPGHHSHGHHPHGHHPHSHSHSGHHPPGHHHPHGHHHPHGHHHPHGHHHPHCH 417

QY 61 DFQDYGPCDPPPHNQ--GHCCHGHGPPPGHLRRRPGKGPFPFHCRCQIGSVYRLPPLRK 117
Db 418 DFLDYGPCDPPSNSQELKGQYHRGYGPPPHGHSRKRGPGLFPFHHQQIGYVYRLPPLNI 477

QY 118 GEVLPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 478 GEVLTLPEANFPSPFLPNCNRSLQPEIQPPF 508

RESULT 8
Q99PS8
ID Q99PS8 PRELIMINARY; PRT; 525 AA.
AC Q99PS8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein 1.
GN Name=RNHRG1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AB055895; BAB33092.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59049 MW; 38290A631FAC7777 CRC64;

Query Match 66.7%; Score 644; DB 2; Length 525;
Best Local Similarity 68.6%; Pred. No. 1.6e-42;
Matches 107; Conservative 7; Mismatches 34; Indels 8; Gaps 2;

QY 1 HPHKHHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHHPHCH 55
Db 342 HPHGHHPHGQHPHGHPHGQHPHGHHPHGQHPHGHHPHGQHPHGHHPHGHDHPHCH 401

QY 56 HPHCHDFQDYGPCDPPPHNQ--GHCCHGHGPPPGHLRRRPGKGPFPFHCRCQIGSVYRL 112
Db 402 HPHGHDFLDYGPCDPPSNSQELKGQYHRGHGPPPHGHSRKRGPGLFPFHHQRIQYVYRL 461

QY 113 PPLRKGEVLPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 462 PPLNVGEVLTPPEANFPFIFSLPNCNRRPPQPEIRPFP 497

RESULT 9
Q9BGU1
ID Q9BGU1 PRELIMINARY; PRT; 546 AA.
AC Q9BGU1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=BTHRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AB055894; BAB33091.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 546 AA; 61948 MW; 26264858824D89EE CRC64;

Query Match 66.7%; Score 644; DB 2; Length 546;
Best Local Similarity 62.7%; Pred. No. 1.7e-42;
Matches 104; Conservative 10; Mismatches 34; Indels 18; Gaps 2;

QY 1 HPHKHHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHPHGHHPHGHHHPHCH 60
Db 353 HPHGHHPHGHPHGHPHGHPHGQHPHGHPHGHPHGHPHGHPHGHPHGHPHCH 412

QY 61 -----DFQDYGPCDPPPHN---QGHCCHGHGPPPGHLRRRPGKGPFPF 102
Db 413 HPGHHPHGHPHDNDFYDHGPCDPPPHRQDPQDHHRQGRDPPPRHSKRGPGKGHFRFY 472

QY 103 CRQIGSVYRLPPLRKGEVLPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 473 WRPTGYIHLPLSLKKGEVLPLPEANFPSPFLPNNNNPLQPEIQAFP 518

RESULT 10
Q99PS7
ID Q99PS7 PRELIMINARY; PRT; 515 AA.
AC Q99PS7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	
DE	Histidine-rich glycoprotein 2.	
GN	Name=RNHRG2;	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Sprague-Dawley; TISSUE=Liver;	
RA	Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,	
RA	Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;	
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB055896; BAB33093.1; -.	
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.	
DR	InterPro; IPR000010; Prot_inh_cystat.	
DR	Pfam; PF00031; Cystatin; 1.	
DR	SMART; SM00043; CY; 2.	
SQ	SEQUENCE 515 AA; 58055 MW; 7CEBA3A1A3678966 CRC64;	
Query Match 64.8%; Score 626; DB 2; Length 515;		
Best Local Similarity 69.5%; Pred. No. 4e-41;		
Matches 105; Conservative 6; Mismatches 32; Indels 8; Gaps 2;		
QY	1 HPHKHSHEQHPGHHPHAHPHEHDTHQHPGHHPHGHHPHGHHPHGHHPHCH 60	
Db	342 HPHGHHPGHQHPGHHPGHQHPGH-----HPHGHHLRHHHPGHHPGHHPGH 396	
QY	61 DFQDYGPCDPPPHNQ---GHCCHGHPPPGHLRRRGKGRPFHCRQIGSVYRLPPLRK 117	
Db	397 DFLDYGPCDPPSNSQDLKGQYHRGHPGHGSRKRGKGLFPFHQRQIGYVYRLPLNV 456	
QY	118 GEVLPLPEANFSPFLPHHKHPLKPDNQPPF 148	
Db	457 GEVLTPEANFSPFSLPNCNRPPQPEIQPPF 487	
RESULT 11		
Q9ESB2	PRELIMINARY; PRT; 510 AA.	
AC	Q9ESB2	
DT	01-MAR-2001 (TrEMBLrel. 16, Created)	
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	
DE	Histidine-rich glycoprotein.	
GN	Name=Hrg;	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Lewis;	
RX	MEDLINE=20307726; PubMed=10849117;	
RA	Hulett M.D., Parish C.R.;	
RT	"Murine histidine-rich glycoprotein: cloning, characterization and cellular origin.";	
RL	Immunol. Cell Biol. 78:280-287(2000).	
DR	EMBL; AF194029; AAG28417.1; -.	
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.	
DR	InterPro; IPR000010; Prot_inh_cystat.	
DR	Pfam; PF00031; Cystatin; 1.	
DR	SMART; SM00043; CY; 2.	
SQ	SEQUENCE 510 AA; 57581 MW; 508E6E06AA2ED58E CRC64;	
Query Match 63.6%; Score 614.5; DB 2; Length 510;		
Best Local Similarity 68.9%; Pred. NO. 3.1e-40;		
Matches 102; Conservative 7; Mismatches 32; Indels 7; Gaps 2;		
QY	1 HPHKHSHEQHPGHHPHAHPHEHDTHQHPGHHPHGHHPHGHHPHGHHPHCH 60	
Db	342 HPHGHHPGHQHPGHHPGHQHPGH-----HPHGHPRGHHHPGHHPGHHPGH 396	

QY	61 DFQDYGPCDPPPHNQCHCHGHPPPGHLRRRGKGRPFHCRQIGSVYRLPPLRKGEV 120	
Db	397 DFLDYGPCDPPSNSQYH--QHGPPPHGHSRKRGPCKGLFPFHQRQIGYVYRLPPLNVGEV 454	
QY	121 LPLPEANFSPFPLPHHKHPLKPDNQPPF 148	
Db	455 LTPPEANFPFISLPNCNRPPQPEIRPPF 482	
RESULT 12		
HRG_RABIT	STANDARD; PRT; 526 AA.	
ID	HRG_RABIT	
AC	Q28640;	
DT	01-NOV-1997 (Rel. 35, Created)	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG) (Fragment).	
DE	Name=HRG;	
GN	Oryctolagus cuniculus (Rabbit).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.	
OX	NCBI_TaxID=9986;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 9-23; 301-313 AND 422-429.	
RC	TISSUE=Serum;	
RX	MEDLINE=96229917; PubMed=8639676; DOI=10.1021/bi952061t;	
RA	Borza D.-B., Tatum F.M., Morgan W.T.;	
RT	"Domain structure and conformation of histidine-proline-rich glycoprotein.";	
RT	Biochemistry 35:1925-1934(1996).	
CC	-!- FUNCTION: The physiological function is not yet known. It binds heme, dyes and divalent metal ions. It can inhibit rosette formation and is known to interact with heparin, thrombospondin, and the lysine-binding site of plasminogen. On the basis of its homology with HMW kininogen, the His-rich region of this protein may mediate the contact activation phase of intrinsic blood coagulation cascade.	
CC	-!- SUBCELLULAR LOCATION: Secreted.	
CC	-!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.	
CC	-!- DOMAIN: In addition to having a high His and Pro content, this protein has many internal repeats. 15 tandem repetitions of a 5-residue sequence (G[H/P]H[P]H, consensus) form a His/Pro-rich region.	
CC	-!- SIMILARITY: Contains 2 cystatin-like domains.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; U32189; AAC48516.1; -.	
DR	InterPro; IPR000010; Prot_inh_cystat.	
DR	Pfam; PF00031; Cystatin; 1.	
KW	Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat; Signal.	
FT	NON_TER 1 1	Potential.
FT	SIGNAL <1 8	Histidine-rich glycoprotein.
FT	CHAIN 9 526	Cystatin-like 1.
FT	DOMAIN 9 126	Cystatin-like 2.
FT	DOMAIN 127 243	Pro-rich.
FT	DOMAIN 251 296	His/Pro-rich.
FT	DOMAIN 329 498	By similarity.
FT	DISULFID 14 505	By similarity.
FT	DISULFID 68 79	By similarity.
FT	DISULFID 95 116	By similarity.
FT	DISULFID 193 415	By similarity.
FT	DISULFID 207 230	By similarity.
FT	DISULFID 272 302	Potential.


```
FT CARBOHYD 115 115 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 192 192 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 240 240 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 310 310 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 485 485 N-linked (GlcNAc. . .) (Potential).
FT SITE 303 304 Cleavage (by plasmin).
FT SITE 421 422 Cleavage (by plasmin).
SQ SEQUENCE 526 AA; 58877 MW; 810F23D367D93D42 CRC64;

Query Match 56.2%; Score 543; DB 1; Length 526;
Best Local Similarity 57.6%; Pred. No. 1.2e-34;
Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

Qy 2 PHKHSHSEQHPHGHPHA-----HH-----PHEHDTHRQHPHGHPHGHP 42
Db 330 PHGHPHGPPPHGHHPHGPPPHGHPHGPPPHGPPPHGHPHGPPPHGHPHGPP 389

Qy 43 HGHHPHGHPHGHPHCHDFQDYGCDPPPHNQG----HCCHGHPGPPPHGLRRRPGKGP 98
Db 390 HGHPHGPPPHGHPHGHPHGFDHGGCDPPSHKEGPQDLH-QHAMGPPPKHPGKRGPGKH 448

Qy 99 RPFHCRQIGSVYRLPPLRKGEVLPPLPEANFPSPFLPHHKHPLKPDNQPPF 148
Db 449 FPFHWRRIQSVYQLPPLQKGEVLPPLPEANFPQLLRNHTHPLKPEIQPPF 498
```

```
RESULT 13
HRG_BOVIN STANDARD; PRT; 396 AA.
ID HRG_BOVIN
AC P33433;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein (Histidine-proline rich glycoprotein)
DE (HPRG) (Fragments).
GN Name=HRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=93351678; PubMed=8348977; DOI=10.1016/0014-5793(93)80945-Q;
RA Soerensen C.B., Krogh-Pedersen H., Petersen T.E.;
RT "Determination of the disulphide bridge arrangement of bovine
RT histidine-rich glycoprotein.";
RL FEBS Lett. 328:285-290(1993).
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 12 tandem repetitions of a 5-
CC residue sequence (GHHPH, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.
DR InterPro; IPR000010; Prot_inh_cystat.
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat.
FT DOMAIN 1 102 Cystatin-like 1.
FT DOMAIN 103 169 Cystatin-like 2.
FT DOMAIN 191 238 Pro-rich.
FT DOMAIN 243 368 His/Pro-rich.
FT DISULFID 7 375
FT DISULFID 56 67
FT DISULFID 77 92
FT DISULFID 123 297
FT DISULFID 137 160
FT DISULFID 212 242
```

```
FT NON CONS 52 53 N-linked (GlcNAc. . .).
FT CARBOHYD 70 70
FT NON CONS 71 72
FT NON CONS 78 79
FT CARBOHYD 91 91 N-linked (GlcNAc. . .).
FT NON CONS 103 104
FT CARBOHYD 122 122 N-linked (GlcNAc. . .).
FT NON CONS 163 164
FT CARBOHYD 220 220 N-linked (GlcNAc. . .).
FT NON CONS 263 264
FT NON CONS 303 304
FT VARIANT 86 86 S -> R.
FT VARIANT 309 309 S -> Q.
FT VARIANT 322 322 H -> Y.
SQ SEQUENCE 396 AA; 44470 MW; 128A8223499DE6FC CRC64;

Query Match 50.6%; Score 489; DB 1; Length 396;
Best Local Similarity 47.8%; Pred. No. 1.5e-30;
Matches 88; Conservative 16; Mismatches 26; Indels 54; Gaps 6;

Qy 1 HPHKHHS-----HEQH-----PH-----GHPHHAHPHE 24
Db 203 HPESYNFRCPPLLEHKNSDSPFQARAPLPFPPLGLRCPHPFPGTKGNHRP----PHD 258

Qy 25 HDTHRQHPHGHPHGHPHGHPHGHPHGHPHCHDFQDYGCDPPPHNQHCCHGHP 84
Db 259 HSSDE-----HHPHGHPHGHPHGHPHGHPHGHPDNDFYDHGPCDPPPHR-----P 304

Qy 85 PPGHLRRRGPGKGRPFHCRQIGSVYRLPPLRKGEVLPPLPEANFPSPFLPHHKHPLKPDN 144
Db 305 PPRHSKERGPGKGHFRFHWRTGYIHRPLPSLKKGEVLPPLPEANFPSPFLPHNHNPLQPEI 364

Qy 145 QPFP 148
Db 365 QAFP 368
```

```
RESULT 14
Q8MP30 PRELIMINARY; PRT; 233 AA.
ID Q8MP30; Q8T164;
AC Q8MP30; Q8T164;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Similar to Plasmodium lophurae. Histidine-rich
DE glycoprotein).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AC123513; AAM44363.1; -.
DR EMBL; AC117070; AAM09303.2; -.
KW Hypothetical protein.
```

SQ SEQUENCE 233 AA; 28937 MW; AD4C7A2F86BB14B9 CRC64;
Query Match 29.1%; Score 281.5; DB 2; Length 233;
Best Local Similarity 38.3%; Pred. No. 1.5e-14;
Matches 57; Conservative 3; Mismatches 62; Indels 27; Gaps 5;
Qy 1 HPHKHSHEQHPHGHHPHAHHPHEHDTHRQHPHGHHHPHGHHHPHGHHHPHGHH 60
Db 79 HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 138
Qy 61 DFQDYGPCDPPPHNQGHCHGHPGPPPGHLRRRGKGPFPFHCRCQIGSVYRLPPLRKGEV 120
Db 139 HHH-----HHHHHHHHHHHHHPHHPHHPHHPHHP-----HLHP----- 173
Qy 121 LPLPEANFPSPFLPH-HKHPLKPDNQFPF 148
Db 174 NPHPHPHPHPHPHPHHHHP-NPNPHPH 201

RESULT 15
HRPX PLALO
ID_HRPX_PLALO STANDARD; PRT; 351 AA.
AC P04929;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Histidine-rich glycoprotein precursor.
OS Plasmodium lophurae.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5853;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061618; PubMed=6095114;
RA Ravetch J.V., Feder R., Pavlovic A., Blobel G.;
RT "Primary structure and genomic organization of the histidine-rich
RT protein of the malaria parasite Plasmodium lophurae.";
RL Nature 312:616-620(1984).
CC -!- MISCELLANEOUS: In the intraerythrocytic stages of development of
CC P.lophurae in ducks, there is a synthesis of a major protein that
CC accumulates to comprise at least 50% of the cellular mass: the
CC histidine rich protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01469; CAA25698.1; -.
DR PIR; A22692; KGZQHL.
DR HSSP; P13231; IHCE.
KW Glycoprotein; Malaria; Repeat; Signal.
FT SIGNAL 1 23
FT PROPEP 24 47
FT CHAIN 48 351
FT CARBOHYD 40 40
FT DOMAIN 59 90
FT REPEAT 59 74
FT REPEAT 75 90
FT DOMAIN 91 123
FT REPEAT 91 107
FT REPEAT 108 123
FT REPEAT 124 153
FT DOMAIN 124 138
FT REPEAT 139 153
FT DOMAIN 173 351
SQ SEQUENCE 351 AA; 44032 MW; D19A48D47D890453 CRC64;

Query Match 25.0%; Score 241.5; DB 1; Length 351;
Best Local Similarity 47.7%; Pred. No. 2.8e-11;
Matches 42; Conservative 1; Mismatches 34; Indels 11; Gaps 3;

Qy 2 PHKHSHSQHPHGHH-----PHAHHPHEH--DTHRQHPHGHHHPHGHHHPHGHHHPHG 54
Db 184 PHHHHHHHHAPHHHHHHHHHAPHHHHHHHHHAPHHHHHHHHHAPHHHHHHHHHAPHHHHHH 243
Qy 55 HHPHCHDFQDYGPCDPPPHNQGHCHG 82
Db 244 HHHHHHHHHHD-----AHHHHHHHHHDANH 267

Search completed: June 15, 2005, 13:51:01
Job time : 135.141 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:31:29 ; Search time 95.7269 Seconds
(without alignments)
408.065 Million cell updates/sec

Title: US-10-074-225A-6
Perfect score: 697
Sequence: 1 SVNIIHRPPPHGHGHPGPP.....HPPHGHGFHDHGPCDPPSHK 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	697	100.0	101	8	ADH10412	Adh10412 Rabbit HP
2	697	100.0	526	5	ABB79805	Abb79805 Rabbit hi
3	697	100.0	526	8	ADH10410	Adh10410 Rabbit HP
4	314	45.1	525	5	ABB79804	Abb79804 Human his
5	314	45.1	525	8	ADE76897	Ade76897 Human pro
6	314	45.1	525	8	ADH10409	Adh10409 Human HPR
7	314	45.1	525	8	ADL12335	Adl12335 Human ste
8	280.5	40.2	148	5	ABB79807	Abb79807 Rabbit hi
9	280.5	40.2	148	5	ABB79806	Abb79806 Human his
10	280.5	40.2	148	8	ADH10411	Adh10411 Human HPR
11	264.5	37.9	491	4	ABG09949	Abg09949 Novel hum
12	259.5	37.2	296	7	ADD45442	Add45442 Rat Prote
13	259.5	37.2	296	7	ADE57105	Ade57105 Rat Prote
14	246	35.3	274	7	ADD47240	Add47240 Rat Prote
15	246	35.3	274	7	ADD48703	Add48703 Rat Prote
16	233.5	33.5	274	2	AAR04119	Aar04119 ORF2 of E
17	233.5	33.5	309	2	AAR04116	Aar04116 ORF1 of E
18	233.5	33.5	331	7	ADD45740	Add45740 Human Pro
19	233.5	33.5	424	6	ABO52997	Abo52997 Human spl
20	233.5	33.5	424	8	ABM80186	Abm80186 Tumour-as
21	231.5	33.2	192	6	AAO30246	Aao30246 Human Sap
22	231.5	33.2	208	6	AAO30188	Aao30188 Human cp3
23	231	33.1	173	7	ADJ68263	Adj68263 Human hea
24	231	33.1	267	7	AAO30408	Aao30408 Human sec
25	229.5	32.9	124	2	AAW50192	Aaw50192 Amino aci

26	229	32.9	230	7	ADG15008	Adg15008 Human SEC
27	228.5	32.8	93	4	AAB71667	Aab71667 Human col
28	228.5	32.8	93	4	AAB71654	Aab71654 Human col
29	228.5	32.8	102	4	ABB68218	Abb68218 Drosophil
30	228	32.7	247	6	ABU11891	Abu11891 Human ABC
31	225	32.3	80	4	ABB61183	Abb61183 Drosophil
32	223	32.0	594	4	ABB61362	Abb61362 Drosophil
33	223	32.0	1130	5	ABP73921	Abp73921 Candida a
34	222	31.9	272	7	ADG15007	Adg15007 Human SEC
35	221	31.7	1378	5	ABB57176	Abb57176 Mouse isc
36	220	31.6	220	8	ADN43006	Adn43006 Human sec
37	215.5	30.9	104	5	ABB78538	Abb78538 Ser-Pro-P
38	215	30.8	162	7	AAO30407	Aao30407 Human sec
39	214.5	30.8	205	8	ADN43007	Adn43007 Human sec
40	213.5	30.6	449	8	ADS23375	Ads23375 Bacterial
41	211	30.3	518	8	ADN20587	Adn20587 Bacterial
42	210	30.1	487	2	AAR22380	Aar22380 Antigen m
43	209.5	30.1	205	5	ABR39650	Abr39650 A. thalia
44	208.5	29.9	163	8	ADM80808	Adm80808 Human CAD
45	208	29.8	865	3	AAB26153	Aab26153 Neospora

ALIGNMENTS

RESULT 1
ADH10412
ID ADH10412 standard; protein; 101 AA.
XX
AC ADH10412;
XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein H/P rich domain.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Oryctolagus cuniculus.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX
DR WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Claim 6; SEQ ID NO 24; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the

CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 697; DB 8; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVNIIHRPPPHGHGPPPPHGHHPHGGPPPHGHPHGGPPPHGHPHGGPPPHGHPHGGPPH 60
Db 1 SVNIIHRPPPHGHGPPPPHGHHPHGGPPPHGHPHGGPPPHGHPHGGPPPHGHPHGGPPH 60
Qy 61 GHPPHGGPPPHGHGPPPPHGHHPHGGPPPHGHGHPHGHGHPHGHGHPHGHGHPHGHGHPH 101
Db 61 GHPPHGGPPPHGHGPPPPHGHHPHGGPPPHGHGHPHGHGHPHGHGHPHGHGHPHGHGHPH 101

RESULT 2
ABB79805
ID ABB79805 standard; protein; 526 AA.

XX ABB79805;

XX 25-NOV-2002 (first entry)

XX Rabbit histidine proline rich glycoprotein.

XX Histidine proline rich glycoprotein; HPRG; rabbit; antiangiogenic;
KW cytosolic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.

OS Oryctolagus cuniculus.

XX
FH Key Location/Qualifiers
FT Domain 251..296
FT /note= "Proline-rich domain"
FT Domain 321..421
FT /note= "Histidine-proline-rich domain, region also
FT specifically claimed in Claim 1"

PN WO200264621-A2.

XX 22-AUG-2002.

PF 14-FEB-2002; 2002WO-US004336.

XX 14-FEB-2001; 2001US-0268370P.

XX (ATTE-) ATTENUON LLC.

XX Donate F, Harris S, Plunkett ML, Mazar AP;

XX WPI; 2002-666989/71.

DR P-PSDB; ABN84911.

XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for

PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 12-13; 82pp; English.

XX The present sequence is the protein sequence of rabbit histidine proline
CC rich glycoprotein (HPRG), a proteinaceous chelator that can be used to
CC inhibit angiogenesis and treat cancer. Claimed anti-angiogenic
CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)
CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see
CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,
CC endothelial cell proliferation or endothelial tube formation in vitro or
CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.
CC Also claimed are: chemically synthesised or recombinantly produced
CC peptide multimers; a diagnostically or therapeutically labeled anti-
CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
CC useful HPRG-related composition, comprising the diagnostically labeled
CC polypeptide, peptide or peptide multimer and a carrier; an antibody
CC specific for an epitope of HPRG that is present in the H/P domain of
CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,
CC or an antigen-binding fragment of the antibody; a method for inhibiting
CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
CC inducing apoptosis; a method for treating a subject having a disease or
CC condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis; a method of stimulating or inhibiting
CC angiogenesis in a subject; a method of detecting the presence of HPRG or
CC its cleavage product or its peptide in a biological sample; isolated
CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an
CC expression vector; transformed or transfected cells; a method of
CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of
CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;
CC an affinity ligand useful for binding to, or isolating, an HPRG-binding
CC molecule or cells expressing the binding molecule, comprising the
CC polypeptide, peptide or peptide multimer, immobilised to a solid support
CC or carrier; and a method of isolating HPRG-binding molecule, or isolating
CC or enriching cells expressing HPRG-binding site or receptor, from a
CC complex mixture. The compositions and methods are useful in diagnosing or
CC treating a disease or condition associated with undesired cell migration,
CC invasion, proliferation, or angiogenesis, such as cancer,
CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,
CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in
CC inhibiting the growth of primary tumours or metastases, and may also be
CC used in treating neurodegenerative diseases like Alzheimer's disease,
CC Parkinson's disease and amyotrophic lateral sclerosis. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX
SQ Sequence 526 AA;

Query Match 100.0%; Score 697; DB 5; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.5e-51;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVNIIHRPPPHGHGPPPPHGHHPHGGPPPHGHPHGGPPPHGHPHGGPPH 60
Db 321 SVNIIHRPPPHGHGPPPPHGHHPHGGPPPHGHPHGGPPPHGHPHGGPPH 380
Qy 61 GHPPHGGPPPHGHGPPPPHGHHPHGGPPPHGHGHPHGHGHPHGHGHPHGHGHPHGHGHPH 101
Db 381 GHPPHGGPPPHGHGPPPPHGHHPHGGPPPHGHGHPHGHGHPHGHGHPHGHGHPHGHGHPH 421

RESULT 3
ADH10410
ID ADH10410 standard; protein; 526 AA.
XX
AC ADH10410;
XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein sequence.

XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT Domain 251..296
FT /note = pro-rich domain
FT Domain 321..421
FT /note = His-Pro rich domain
XX
PN WO2003077872-A2.
XX
XX 25-SEP-2003.
XX
XX 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
XX (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
XX
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Claim 6; SEQ ID NO 22; 117pp; English.
XX
XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
XX protein sequence.
SQ Sequence 526 AA;

Query Match 100.0%; Score 697; DB 8; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.5e-51;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVNIIHRPPPHGHPHPPHGHPPHPPHGHPPHPPHPPHPPHPPHPPHPPHPPH 60
|||||
Db 321 SVNIIHRPPPHGHPHPPHGHPPHPPHPPHPPHPPHPPHPPHPPH 380
|||||

Qy 61 GHPPHGGPPPHGHPHPPHPPHGHPPHPPHGHPPHPPHGHPPHPPH 101
|||||
Db 381 GHPPHGGPPPHGHPHPPHPPHGHPPHPPHGHPPHPPHGHPPHPPH 421
|||||

RESULT 4
ABB79804
ID ABB79804 standard; protein; 525 AA.
XX
AC ABB79804;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human histidine proline rich glycoprotein.
XX
KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 276..321
FT /note= "Proline-rich domain"
FT Domain 350..497
FT /note= "Histidine-proline-rich domain, region also
FT specifically claimed in Claim 1"
XX
PN WO200264621-A2.
XX
XX 22-AUG-2002.
XX
XX 14-FEB-2002; 2002WO-US004336.
XX
XX 14-FEB-2001; 2001US-0268370P.
XX (ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 11; 82pp; English.
XX
XX The present sequence is the protein sequence of human histidine proline
CC rich glycoprotein (HPRG), a proteinaceous metal chelator that can be used
CC to inhibit angiogenesis and treat cancer. Claimed anti-angiogenic
CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)
CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see
CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,
CC endothelial cell proliferation or endothelial tube formation in vitro or
CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.
CC Also claimed are: chemically synthesised or recombinantly produced
CC peptide multimers; a diagnostically or therapeutically labeled anti-
CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
CC useful HPRG-related composition, comprising the diagnostically labeled
CC polypeptide, peptide or peptide multimer and a carrier; an antibody
CC specific for an epitope of HPRG that is present in the H/P domain of
CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,
CC or an antigen-binding fragment of the antibody; a method for inhibiting
CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
CC inducing apoptosis; a method for treating a subject having a disease or
CC condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis; a method of stimulating or inhibiting
CC angiogenesis in a subject; a method of detecting the presence of HPRG or
CC its cleavage product or its peptide in a biological sample; isolated

CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an
CC expression vector; transfected or transfected cells; a method of
CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of
CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;
CC an affinity ligand useful for binding to, or isolating, an HPRG-binding
CC molecule or cells expressing the binding molecule, comprising the
CC polypeptide, peptide or peptide multimer, immobilised to a solid support
CC or carrier; and a method of isolating HPRG-binding molecule, or isolating
CC or enriching cells expressing HPRG-binding site or receptor, from a
CC complex mixture. The compositions and methods are useful in diagnosing or
CC treating a disease or condition associated with undesired cell migration,
CC invasion, proliferation, or angiogenesis, such as cancer,
CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,
CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in
CC inhibiting the growth of primary tumours or metastases, and may also be
CC used in treating neurodegenerative diseases like Alzheimer's disease,
CC Parkinson's disease, and amyotrophic lateral sclerosis. The antibodies
CC are stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 525 AA;

Query Match 45.1%; Score 314; DB 5; Length 525;
Best Local Similarity 41.8%; Pred. No. 9.9e-19;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;

Qy 7 RPP--PHG-----HHPHGPPPHGHHPHGPPP-----HGHP-PHGPPP----- 40
Db 275 KPFFKPHGSRDHHHPHKP-----HEHGPPPPPPDERDHSHGPPPLQGPPPLLPMSCSSCQH 329

Qy 41 -----RHP-----RHPPPHGHHP-----HGPPPHGHHP----- 67
Db 330 ATFGTNGAQRHSHNNSSDLHPHKHSHSQHPHGHHPHAHPHEHDTHRQHPHGHHPHGH 389

Qy 68 PPHGHPPHGGPPPHGHPPHGHGHFDHGHGCDPPSH 100
Db 390 HPHGHHPHGHHPHGHHPHCHDFQDYGPCDPPPH 422

RESULT 5
ADE76897
ID ADE76897 standard; protein; 525 AA.
XX
AC ADE76897;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein expressed in a liver disorder #18.
XX
KW human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003108871-A1.
XX
PD 12-JUN-2003.
XX
PF 30-JUL-2001; 2001US-00919039.
XX
PR 28-JUL-2000; 2000US-0222113P.
XX
PA (KASE/) KASER M R.
XX
PI Kaser MR;
XX
DR WPI; 2004-031227/03.
DR N-PSDB; ADE76896.
XX
PT Composition comprising several cDNAs that are differentially expressed in
treated human C3A liver cell cultures, useful for treating liver

PT disorders.
XX Claim 1; SEQ ID NO 62; 41pp; English.
XX
CC The invention relates to a composition comprising several cDNAs that are
CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.
XX
SQ Sequence 525 AA;

Query Match 45.1%; Score 314; DB 8; Length 525;
Best Local Similarity 41.8%; Pred. No. 9.9e-19;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;

Qy 7 RPP--PHG-----HHPHGPPPHGHHPHGPPP-----HGHP-PHGPPP----- 40
Db 275 KPFFKPHGSRDHHHPHKP-----HEHGPPPPPPDERDHSHGPPPLQGPPPLLPMSCSSCQH 329

Qy 41 -----RHP-----RHPPPHGHHP-----HGPPPHGHHP----- 67
Db 330 ATFGTNGAQRHSHNNSSDLHPHKHSHSQHPHGHHPHAHPHEHDTHRQHPHGHHPHGH 389

Qy 68 PPHGHPPHGGPPPHGHPPHGHGHFDHGHGCDPPSH 100
Db 390 HPHGHHPHGHHPHGHHPHCHDFQDYGPCDPPPH 422

RESULT 6
ADH10409
ID ADH10409 standard; protein; 525 AA.
XX
AC ADH10409;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human HPRG protein sequence.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
Domain 276..321
FT /note = pro-rich domain
FT Domain 350..497
FT /note = His-Pro rich domain
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX

PR 15-MAR-2002; 2002US-0364047P.
XX (ATTE-) ATTENUON LLC.
PA
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX
XX WPI; 2004-090604/09.
DR
XX
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
XX Claim 6; SEQ ID NO 21; 117pp; English.
PS
XX
XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a human HPRG
CC protein sequence.
XX
SQ Sequence 525 AA;

Query Match 45.1%; Score 314; DB 8; Length 525;
Best Local Similarity 41.8%; Pred. No. 9.9e-19;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;
QY 7 RPP--PHG---HHPHGPPPHGHHPHGPPP-----HGHP-PHGPPP----- 40
Db 275 KPPEKPHGSRDHHHPK-----HEHGPPPPDERDHSHPPLPQGPPPLPMSCSSCQH 329
QY 41 -----RHP-----PHGPPPHGHPP-----HGPPPHGHPPHPG 67
Db 330 ATFGTNGAQRSHNNSSDLHPKHHSHEQHPGHHPHAHPHEHDTHQHPGHHPHGH 389
QY 68 PPHGHPPPHGPPPHGHHPGHGFDHGCDPPSH 100
Db 390 HPHGHHPGHHPGHHPGHCHDFQDYGCDPPPH 422

RESULT 7
ADL12335
ID ADL12335 standard; protein; 525 AA.
XX
AC ADL12335;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell protein #11.
XX
KW Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.

XX Homo sapiens.
OS
XX
PN US6673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Furness LM, Buchbinder JL;
XX
DR WPI; 2004-068610/07.
XX
PT Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
PS Disclosure; SEQ ID NO 64; 141pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human protein which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 525 AA;

Query Match 45.1%; Score 314; DB 8; Length 525;
Best Local Similarity 41.8%; Pred. No. 9.9e-19;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;
QY 7 RPP--PHG---HHPHGPPPHGHHPHGPPP-----HGHP-PHGPPP----- 40
Db 275 KPPEKPHGSRDHHHPK-----HEHGPPPPDERDHSHPPLPQGPPPLPMSCSSCQH 329
QY 41 -----RHP-----PHGPPPHGHPP-----HGPPPHGHPPHPG 67
Db 330 ATFGTNGAQRSHNNSSDLHPKHHSHEQHPGHHPHAHPHEHDTHQHPGHHPHGH 389
QY 68 PPHGHPPPHGPPPHGHHPGHGFDHGCDPPSH 100
Db 390 HPHGHHPGHHPGHHPGHCHDFQDYGCDPPPH 422

RESULT 8
ABB79807
ID ABB79807 standard; protein; 148 AA.
XX
AC ABB79807;
XX
DT 25-NOV-2002 (first entry)
XX
DE Rabbit histidine proline rich glycoprotein His/Pro-rich domain.
XX

KW Histidine proline rich glycoprotein; HPGP; rabbit; antiangiogenic;
KW cyostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
XX
OS Oryctolagus cuniculus.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 13; 82pp; English.
XX
CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of rabbit anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79805). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain (see ABB79806) of human
CC HPRG; an H/P domain of rabbit HPRG; a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier, and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 148 AA;

Query Match 40.2%; Score 280.5; DB 5; Length 148;
Best Local Similarity 51.6%; Pred No. 2.3e-16;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;
QY 10 PHGHPHGGPPPHGHPHGGPPPHGHPHGGPPPHGHPHGGPPPHGHPHGGPP 69
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 PHKHSHEQHPHGHPHA-----HH-----PHEHDTHRQHPHGHPHGHP 42
QY 70 HGHPPHGGPPPHGHPHGGHPHGHGFDHGDGCDPPSH 100
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 43 HGHHPHGHHPHGHPHCHDFQDYGCDPPPH 73
RESULT 9
ABB79806
ID ABB79806 standard; protein; 148 AA.
XX
AC ABB79806;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human histidine proline rich glycoprotein His/Pro-rich domain.
XX
KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cyostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian.
XX
OS Homo sapiens.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX (ATTE-) ATTENUON LLC.
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 13; 82pp; English.
XX
CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of human anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79804). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain of human HPRG; an H/P
CC domain of rabbit HPRG (see ABB12345); a variant of these that is capable
CC of inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier, and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 148 AA;

CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transduced or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 148 AA;

Query Match 40.2%; Score 280.5; DB 5; Length 148;
Best Local Similarity 51.6%; Pred. No. 2.3e-16;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;

Qy 10 PHGHHPPPPHGHHPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPH 69
Db 2 PHKHSHEQHHPHGHHPH-----HH-----PHEHDTHRQHPHGHHPHGHHP 42

Qy 70 HGHPPHPPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPH 100
Db 43 HGHHPHGHHPHGHHPHCHDFQDYGCDPPPH 73

RESULT 10
ADH10411
ID ADH10411 standard; protein; 148 AA.
XX
AC ADH10411;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human HPRG protein H/P rich domain.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritis;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX
DR WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX

PS Claim 6; SEQ ID NO 23; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a human HPRG
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 148 AA;

Query Match 40.2%; Score 280.5; DB 8; Length 148;
Best Local Similarity 51.6%; Pred. No. 2.3e-16;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;

Qy 10 PHGHHPPPPHGHHPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPH 69
Db 2 PHKHSHEQHHPHGHHPH-----HH-----PHEHDTHRQHPHGHHPHGHHP 42

Qy 70 HGHPPHPPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPHGHPPH 100
Db 43 HGHHPHGHHPHGHHPHCHDFQDYGCDPPPH 73

RESULT 11
ABG09949
ID ABG09949 standard; protein; 491 AA.
XX
AC ABG09949;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9940.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.

DR N-PSDB; AAS74136.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 40308; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have application in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 491 AA;
Query Match 37.9%; Score 264.5; DB 4; Length 491;
Best Local Similarity 41.8%; Pred. No. 1.3e-14;
Matches 56; Conservative 2; Mismatches 35; Indels 41; Gaps 8;

Qy 6 HRPPP-----HGHPHG-----PPPHGHPH-GPPPHGHPHPGPPRHPPH----- 45
Dbb 136 HRPPAQTSTGHHQHGHQHRPPPAHGHQHQSPPAHGHQHQRPPAHGHQHGHQHQS 195
Qy 46 GPPPHGHPHPGPPPHGHPGPP-----PHGHPHPGPP-----HGH-----PPHG 86
Dbb 196 SPPAHGHQHRRHQPCHHQRSPHSTAHHQHGHQHSSPPAHGHQHGHQHSSPYQ 255

Qy 87 HGFHDHGPCDPPSH 100
Dbb 256 HGHQHHRP--PPAH 267

RESULT 12
ADD45442
ID ADD45442 standard; protein; 296 AA.
XX
AC ADD45442;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein L17318, SEQ ID NO 10875.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; L17318.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 296 AA;

Query Match 37.2%; Score 259.5; DB 7; Length 296;
Best Local Similarity 38.4%; Pred. No. 2.3e-14;
Matches 63; Conservative 1; Mismatches 27; Indels 73; Gaps 10;

Qy 2 VNIHRRPPHG---HHPH-----GPPPHG-----HHPHGPPPHGHP-- 34
Dbb 81 VNRPERPPQHGGNHHHPHPPAAGQRPPQPGSPQGGPPPPGGPQORPPQGGPQQR 140
Qy 35 -----PHGPPP-----RHPPHGGPPPHGHP-----PHGPPPHGHP 71
Dbb 141 PPQPGSPQGGPPPPGGPQORPPQGGPPPPQGGPQORPPQPGSPQGGPQORAPQGGPPQ 200
Qy 72 HP-----PHGP-----PPHGHPPHGHGFHDHGPCDPP 98
Dbb 201 GPQRPPQPGSPQGGPPPPGGPQORPPQGGPPQ-----GPQRPP 238

RESULT 13
ADE57105
ID ADE57105 standard; protein; 296 AA.
XX
AC ADE57105;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein L17318, SEQ ID NO 2965.
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:38:35 ; Search time 25.1486 Seconds
(without alignments)
299.800 Million cell updates/sec

Title: US-10-074-225A-6
Perfect score: 697
Sequence: 1 SVNIIHRPPPHGHPHGPPP.....HPPHGHGFHDHGPCDPPSHK 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10
                  Listing first 45

```

```
Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #			ID	Description
	Score	Match	Length		
1	314	45.1	525	US-09-976-594-64	Sequence 64, Appl
2	314	45.1	525	US-09-919-039-62	Sequence 62, Appl
3	233.5	33.5	331	US-09-538-092-845	Sequence 845, App
4	233.5	33.5	424	US-09-538-092-1338	Sequence 1338, Ap
5	231	33.1	320	US-09-949-016-8196	Sequence 8196, Ap
6	229.5	32.9	124	US-08-925-237-2	Sequence 2, Appli
7	229.5	32.9	251	US-09-538-092-840	Sequence 840, App
8	228	32.7	247	US-09-538-092-890	Sequence 890, App
9	224	32.1	276	US-09-538-092-889	Sequence 889, App
10	221.5	31.8	234	US-09-538-092-888	Sequence 888, App
11	220	31.6	258	US-09-949-016-9409	Sequence 9409, Ap
12	215.5	30.9	104	US-09-547-693-235	Sequence 235, App
13	208	29.8	865	US-09-281-766-19	Sequence 19, Appl
14	208	29.8	865	US-09-612-858-19	Sequence 19, Appl
15	208	29.8	865	US-09-957-995A-19	Sequence 19, Appl
16	208	29.8	971	US-09-248-796A-19531	Sequence 19531, A
17	205	29.4	130	US-09-270-767-39206	Sequence 39206, A
18	205	29.4	130	US-09-270-767-54423	Sequence 54423, A
19	196	28.1	82	US-08-925-237-4	Sequence 4, Appli
20	192.5	27.6	511	US-09-107-433-2904	Sequence 2904, Ap
21	192.5	27.6	955	US-09-949-016-8369	Sequence 8369, Ap
22	192	27.5	324	US-09-949-016-7664	Sequence 7664, Ap
23	191.5	27.5	553	US-09-949-016-7961	Sequence 7961, Ap
24	190	27.3	166	US-09-949-016-6709	Sequence 6709, Ap
25	190	27.3	178	US-09-949-016-8394	Sequence 8394, Ap
26	175	25.1	306	US-08-217-327-6	Sequence 6, Appli
27	175	25.1	623	US-09-949-016-6530	Sequence 6530, Ap

ALIGNMENTS

RESULT 1

```

US-09-976-594-64
; Sequence 64, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/9
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,40
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6
US-09-976-594-64

```

Query Match	45.1%;	Score 314;	DB 4;	Length 525;	
Best Local Similarity	41.8%;	Pred: No. 4.6e-21;			
Matches	64;	Conservative	2;	Mismatches 23; Indels 64; Gaps 8;	
Qy	7	RPP--PHG----	HHPHGPPPHGHHPHGPPP-----	HGHP-PHGPPP-----	40
		:			
Db	275	KPFFKPGSRDHHHPKP-----	HEHGPPPPPPDERDHSHGPPPLPQGPPPLLPMSCSSCQH		329
Qy	41	-----RHP-----	PHGPPPHGHP-----	HGPPPHGHPHPGP	67
Db	330	ATFGTNGAQRSSNNSSDLHPHKHSHSEQPHGHHPHAHHPHENDHTRQHPHGHHPHGH			389
Qy	68	PPHGHPPHGGPPPHGHPPHGHGFHDHGPCDPPSH			100
				:	
Db	390	HPHGHHPHGHHPHGHHPHGHCHDFODYGPCDPPPH			422

RESULT 2

US-09-919-039-62
; Sequence 62, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919.039

Db 166 PPHGKPEGPPPPQGGNQSQGPPRRPGKPEGPPPPQGGNQSQGPPRRPGKPEGPPPPQGGNQ 225
Qy 51 -----GHPHGGPPPHGHPHGGPPP-HGHPPHGGPPPHGHP--PHGHGF 89
Db 226 SQGPPRRPGKPEGSPSQGGNKPQGGPPPHGKPGQGGPPPPQEGNKPQRPPPPGRPQGGPPPGG 285
Qy 90 HDHGPCDPPSHK 101
Db 286 NPQQFLPPAGK 297

RESULT 6
US-08-925-237-2
; Sequence 2, Application US/08925237
; Patent No. 5981720
; GENERAL INFORMATION:
; APPLICANT: Azen, Edwin A.
; APPLICANT: Pan, David
; TITLE OF INVENTION: Human Salivary Proteins And Fragments
; TITLE OF INVENTION: Thereof Having Alpha-Glucosidase Inhibitory Activity
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kent Barta
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: WI
; COUNTRY: USA
; ZIP: 53202-4108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,237
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Barta, Kent S.
; REGISTRATION NUMBER: 29,042
; REFERENCE/DOCKET NUMBER: 96429/9002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 283-2275
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-925-237-2

Query Match 32.9%; Score 229.5; DB 2; Length 124;
Best Local Similarity 44.9%; Pred. No. 4.8e-14;
Matches 48; Conservative 3; Mismatches 35; Indels 21; Gaps 5;
Qy 8 PPHGHHPHGGPPPHGHHPHGGPPH-GHPPHGGPPRRHPPHGGPPHGH-----PPH 55
Db 10 PQGGNQPGPPPPPGKPGQGGPPPPQGGNKPQGGPPPPGKPGQGGPPPPQGGNKSQSARSPPGKPG 69
Qy 56 GPPPH-GHPPHGGPPPHGHPHGGPPPHGHPHGHGFDHGGPCDPPSHK 101
Db 70 GPPFQGGNQPGPPPPPGKPGQGGPPPPQGGNKSQGP-----GNKSQGP--PPPGK 108

RESULT 7
US-09-538-092-840
; Sequence 840, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 840
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P02812
US-09-538-092-840
Query Match 32.9%; Score 229.5; DB 4; Length 251;
Best Local Similarity 43.6%; Pred. No. 9.1e-14;
Matches 48; Conservative 3; Mismatches 34; Indels 25; Gaps 5;
Qy 8 PPHGHHPHGGPPPHGHHPHGGPPH-GHPPHGGPPRRHPPHGGPPHGH-----PPH 55
Db 19 PQGGNQPGPPPPPGKPGQGGPPPPQGGNKPQGGPPPPGKPGQGGPPPPQGGNKSQSARSPPGKPG 78
Qy 56 GPPPH-GHPPHGGPPPHGHPHGGPPHGH-----PPHGHGFDHGGPCDPP 98
Db 79 GPPFQGGNQPGPPPPPGKPGQGGPPPPQGGNKSQGGPPPPGK-----PQGPP 122

RESULT 8
US-09-538-092-890
; Sequence 890, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 890
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P10163
US-09-538-092-890
Query Match 32.7%; Score 228; DB 4; Length 247;
Best Local Similarity 36.4%; Pred. No. 1.2e-13;
Matches 52; Conservative 4; Mismatches 39; Indels 48; Gaps 5;
Qy 7 RPPHGHHPHGGPPPHGHHPHGGPPHGH-----HPPHGGPPRRHPPHGH 52
Db 81 RPPQGGNQSQGPPPPHGGKPERPPPPQGGNQSQGTTPPPPGKPERPPPPQGGNQSHRPPPPGK 140
Qy 53 P-----PHGPPPHGHPHGGPPPP-----HGHPHGGPPPH 80
Db 141 PERPPPPQGGNQSQGPPPPHGGKPERPPPPQGGNKSRSARSPPGKPGQGGNKPQGGPPPP 200
Qy 81 GHP--PHGHGFDHGGPCDPPSHK 101

Db 201 GKPGPPAGGNPQQPDPPAGK 223

RESULT 9

US-09-538-092-889

; Sequence 889, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387

; SOFTWARE: CuratPatSeqFormatter Version 0.9

; SEQ ID NO 889

; LENGTH: 276

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)...(0)

; OTHER INFORMATION: Polypeptide Accession Number P10162

US-09-538-092-889

Query Match 32.1%; Score 224; DB 4; Length 276;

Best Local Similarity 37.6%; Pred. No. 3.1e-13;

Matches 50; Conservative 3; Mismatches 36; Indels 44; Gaps 5;

Qy 7 RPPPHGHHPHGPPPHGHHPHGGPPPHG-----HPPHGGPPRHPP-----HGPPPHGHP 53

Db 131 RPPQGGNQSQGPPPHGCKPERPPPPQGGNQSHRPPPPPGKPERPPPPQGGNQSQGPPPHPGK 190

Qy 54 PHGPPP-----HGHPHGGPPPHGHHPHGGPPPHGH-----PPH 85

Db 191 PEGPPPEGKNKSARSAPPGKPGQPPPEGKNKPGGPPPPGKPGQPPPPGPNPQQPQAPPA 250

Qy 86 GHGFHDHGCDPP 98

Db 251 G---KPQGPPPPP 260

RESULT 10

US-09-538-092-888

; Sequence 888, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387

; SOFTWARE: CuratPatSeqFormatter Version 0.9

; SEQ ID NO 888

; LENGTH: 234

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)...(0)

; OTHER INFORMATION: Polypeptide Accession Number P10161

US-09-538-092-888

Query Match 31.8%; Score 221.5; DB 4; Length 234;

Best Local Similarity 42.3%; Pred. No. 4.5e-13;

Matches 47; Conservative 5; Mismatches 40; Indels 19; Gaps 5;

Qy 7 RPPPHGHHPHGPPPH-GHHPHGPPPHGHHPHGGPPR--HPPHGGPPPH-----GH 52

Db 15 RPPPPPGKPGQPPPPQGGNQSGPPPHGKPERPPPPQGGNQSQGPPPHGKPERPPPPQGGN 74

Qy 53 PPHGPPPHGHPPHGGPPPH-GHPPHGGPPPH----GHPPHGHGFHDHGCDPP 98

Db 75 QSQGPPPTPGKPEGPPPPQGGNQSQGPPPHGKPERPPPPQGGNQSHRPPPPP 125

RESULT 11

US-09-949-016-9409

; Sequence 9409, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9409

; LENGTH: 258

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-9409

Query Match 31.6%; Score 220; DB 4; Length 258;

Best Local Similarity 35.7%; Pred. No. 6.7e-13;

Matches 51; Conservative 4; Mismatches 40; Indels 48; Gaps 5;

Qy 7 RPPPHGHHPHGPPPHGHHPHGGPPPHG-----HPPHGGPPRHPPHG-----PPPHGH 52

Db 92 RPPQGGNQSQGPPPHGKPERPPPPQGGNQSQGTTPPPGKPERPPPPQGGNQSHRPPPPGK 151

Qy 53 P-----PHGPPPHGHHPHGGPP-----HGHPHGGPPPH 80

Db 152 PERPPPPQGGNQSQGPPPHGKPEGPPPPQEGNKSARSAPPGKPGQPPPEGKNKPGQPPPP 211

Qy 81 GHP--PHGHGFHDHGCDPPPSHK 101

Db 212 GKPGPPPPAGGNPQQPQAPPAGK 234

RESULT 12

US-09-547-693-235

; Sequence 235, Application US/09547693

; Patent No. 6639050

; GENERAL INFORMATION:

; APPLICANT: Kieliszewski, Marcia

; TITLE OF INVENTION: Synthetic Genes for Plant Gums and Other Hydroxyproline-Rich

; TITLE OF INVENTION: Glycoproteins

; FILE REFERENCE: OHU-04089

; CURRENT APPLICATION NUMBER: US/09/547,693

; CURRENT FILING DATE: 2000-04-12

; NUMBER OF SEQ ID NOS: 236

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 235

; LENGTH: 104

; TYPE: PRT

; ORGANISM: Artificial/Unknown

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:51:11 ; Search time 84.7751 Seconds
(without alignments)
456.700 Million cell updates/sec

Title: US-10-074-225A-6
Perfect score: 697
Sequence: 1 SVNIIHRPPPHGHHPGPPP.....HPPHGHGFHDHGCDPPSHK 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	697	100.0	101	14	US-10-074-225A-6
2	697	100.0	526	14	US-10-074-225A-3
3	314	45.1	525	10	US-09-919-039-62
4	314	45.1	525	14	US-10-074-225A-1
5	314	45.1	525	17	US-10-868-577A-42
6	285.5	41.0	371	16	US-10-425-115-320892
7	280.5	40.2	148	14	US-10-074-225A-5
8	248.5	35.7	235	16	US-10-437-963-171176
9	246	35.3	274	9	US-09-850-887-4
10	241.5	34.6	1643	17	US-10-732-923-8217
11	237	34.0	149	16	US-10-425-115-252326

12	233.5	33.5	241	15	US-10-424-599-223180	Sequence 223180,
13	233.5	33.5	309	16	US-10-751-014-2	Sequence 2, Appli
14	233.5	33.5	309	16	US-10-751-014-5	Sequence 5, Appli
15	232.5	33.4	342	15	US-10-424-599-281824	Sequence 281824,
16	231.5	33.2	208	17	US-10-495-300-40	Sequence 40, Appl
17	231	33.1	173	16	US-10-408-765A-69	Sequence 69, Appl
18	231	33.1	578	16	US-10-425-115-256490	Sequence 256490,
19	228.5	32.8	93	10	US-09-997-003-31	Sequence 31, Appl
20	228.5	32.8	93	10	US-09-997-003-44	Sequence 44, Appl
21	227.5	32.6	107	15	US-10-424-599-234553	Sequence 234553,
22	226.5	32.5	280	16	US-10-425-115-231956	Sequence 231956,
23	225	32.3	204	15	US-10-424-599-221495	Sequence 221495,
24	223	32.0	399	15	US-10-424-599-191976	Sequence 191976,
25	223	32.0	1130	14	US-10-032-585-7758	Sequence 7758, Ap
26	222	31.9	137	16	US-10-425-115-292155	Sequence 292155,
27	221.5	31.8	183	15	US-10-424-599-218680	Sequence 218680,
28	221.5	31.8	208	16	US-10-437-963-103915	Sequence 103915,
29	219.5	31.5	123	16	US-10-425-115-316409	Sequence 316409,
30	219	31.4	296	15	US-10-424-599-224246	Sequence 224246,
31	218.5	31.3	147	15	US-10-424-599-154677	Sequence 154677,
32	217.5	31.2	134	15	US-10-424-599-208512	Sequence 208512,
33	215.5	30.9	104	15	US-10-437-708-235	Sequence 235, App
34	215.5	30.9	104	16	US-10-257-199-235	Sequence 235, App
35	215.5	30.9	104	17	US-10-418-032-235	Sequence 235, App
36	215.5	30.9	230	17	US-10-418-032-269	Sequence 269, App
37	215	30.8	179	15	US-10-424-599-218681	Sequence 218681,
38	213.5	30.6	449	15	US-10-369-493-12408	Sequence 12408, A
39	213	30.6	184	15	US-10-424-599-281818	Sequence 281818,
40	212	30.4	163	15	US-10-424-599-281820	Sequence 281820,
41	212	30.4	178	15	US-10-424-599-264218	Sequence 264218,
42	211.5	30.3	241	15	US-10-424-599-163337	Sequence 163337,
43	211	30.3	136	16	US-10-425-115-222187	Sequence 222187,
44	211	30.3	327	16	US-10-425-115-276286	Sequence 276286,
45	211	30.3	518	15	US-10-369-493-3240	Sequence 3240, Ap

ALIGNMENTS

RESULT 1
US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074, 225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match	100.0%;	Score 697;	DB 14;	Length 101;
Best Local Similarity	100.0%;	Pred. No. 1.3e-42;		
Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	SVNIIHRPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPH 60		
Db	1	SVNIIHRPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPH 60		
QY	61	GHPPHGPppHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPH 101		


```

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 223180
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_43562C.1.pep
US-10-424-599-223180

Query Match      33.5%; Score 233.5; DB 15; Length 241;
Best Local Similarity 48.0%; Pred. No. 1.7e-09;
Matches 49; Conservative 3; Mismatches 39; Indels 11; Gaps 7;

Qy  7 RPPPHGHHPPHGGPPPHGHHPPHGGPPPHGHHPPHPPH-GPPPHGHHPPH-GPPPHGHP 63
Db  27 KPPPEYQPPHEKPPH-----ENPPPEHQPPHEKPPPEHQPHEKPPPEYEPHEKPPPEYQP 82

Qy  64 PH-GPPPHGHHPPH-GPPPHGHHPPHGGHGFHDHGPC--DPPSHK 101
Db  83 PHEKPPPEYQPPHEKPPPEYQPPHEKPPPEHQPHEKPPPEHQ 124
```

```

RESULT 13
US-10-751-014-2
; Sequence 2, Application US/10751014
; Publication No. US20040172676A1
; GENERAL INFORMATION:
; APPLICANT: Franssen, Henk J
; APPLICANT: Bisseling, Anton H
; TITLE OF INVENTION: ENOD2 Gene Regulatory Region
; FILE REFERENCE: MPS 4-87FD2
; CURRENT APPLICATION NUMBER: US/10/751,014
; CURRENT FILING DATE: 2004-01-02
; PRIOR APPLICATION NUMBER: US/09/564,142A
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 08/859,555
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Glycine max
US-10-751-014-2
```

```

Query Match      33.5%; Score 233.5; DB 16; Length 309;
Best Local Similarity 48.0%; Pred. No. 2.1e-09;
Matches 49; Conservative 3; Mismatches 39; Indels 11; Gaps 7;

Qy  7 RPPPHGHHPPHGGPPPHGHHPPHGGPPPHGHHPPHPPH-GPPPHGHHPPH-GPPPHGHP 63
Db  95 KPPPEYQPPHEKPPH-----ENPPPEHQPPHEKPPPEHQPHEKPPPEYEPHEKPPPEYQP 150

Qy  64 PH-GPPPHGHHPPH-GPPPHGHHPPHGGHGFHDHGPC--DPPSHK 101
Db  151 PHEKPPPEYQPPHEKPPPEYQPPHEKPPPEHQPHEKPPPEHQ 192
```

```

RESULT 14
US-10-751-014-5
; Sequence 5, Application US/10751014
; Publication No. US20040172676A1
; GENERAL INFORMATION:
; APPLICANT: Franssen, Henk J
; APPLICANT: Bisseling, Anton H
; TITLE OF INVENTION: ENOD2 Gene Regulatory Region
; FILE REFERENCE: MPS 4-87FD2
; CURRENT APPLICATION NUMBER: US/10/751,014
; CURRENT FILING DATE: 2004-01-02
; PRIOR APPLICATION NUMBER: US/09/564,142A
; PRIOR FILING DATE: 2000-05-03
```

```

; PRIOR APPLICATION NUMBER: 08/859,555
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Glycine max
US-10-751-014-5

Query Match      33.5%; Score 233.5; DB 16; Length 309;
Best Local Similarity 48.0%; Pred. No. 2.1e-09;
Matches 49; Conservative 3; Mismatches 39; Indels 11; Gaps 7;

Qy  7 RPPPHGHHPPHGGPPPHGHHPPHGGPPPHGHHPPHPPH-GPPPHGHHPPH-GPPPHGHP 63
Db  95 KPPPEYQPPHEKPPH-----ENPPPEHQPPHEKPPPEHQPHEKPPPEYEPHEKPPPEYQP 150

Qy  64 PH-GPPPHGHHPPH-GPPPHGHHPPHGGHGFHDHGPC--DPPSHK 101
Db  151 PHEKPPPEYQPPHEKPPPEYQPPHEKPPPEHQPHEKPPPEHQ 192
```

```

RESULT 15
US-10-424-599-281824
; Sequence 281824, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 281824
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(342)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_96507C.1.pep
US-10-424-599-281824
```

```

Query Match      33.4%; Score 232.5; DB 15; Length 342;
Best Local Similarity 48.4%; Pred. No. 2.6e-09;
Matches 61; Conservative 1; Mismatches 25; Indels 39; Gaps 16;

Qy  8 PPHGHHPHGGPPPHGH-HPHGPPPHGHPHPPH-HGPPPRHPP-----HGPPP--HGHP-P 54
Db  179 PVPVHPYPH---PHPHPHH-PHPHHPHPPVYVHSPPP--PPKPYVYVHSPPPPVHHHPYP 232

Qy  55 H-GPPPHGHPH-----HGPPP--HGHP-PHGPppPHGHP-PHGHP----C 95
Db  233 HLHPHHPHHPHPPPPKPYVYVHSPPPPVHPHPYPH-PHPHPHPYPHPHPPHPPPYVHS 291

Qy  96 DPPSHK 101
Db  292 PPPPPK 297
```

Search completed: June 15, 2005, 14:09:36
Job time : 85.7751 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:37:20 ; Search time 19.4699 Seconds
(without alignments)
499.124 Million cell updates/sec

Title: US-10-074-225A-6
Perfect score: 697
Sequence: 1 SVNIIHRPPPHGHHPHGPPP.....HPPHGHGFHDHGCDPPSHK 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	314	45.1	525	1	KGHUGH	histidine-rich gly
2	260.5	37.4	445	2	A60488	histidine-rich gly
3	259.5	37.2	295	2	B48013	proline-rich prote
4	250	35.9	128	2	D38355	basic proline-rich
5	248	35.6	278	2	B39066	proline-rich prote
6	247.5	35.5	206	1	PIRT3	acidic proline-ric
7	246.5	35.4	227	2	C29149	proline-rich prote
8	246.5	35.4	240	2	A24264	proline-rich prote
9	246.5	35.4	317	2	A28996	proline-rich prote
10	239	34.3	240	2	B24264	proline-rich prote
11	236	33.9	330	2	S22140	nodulin Enod2 - Se
12	235.5	33.8	392	1	PIHUB6	salivary proline-r
13	235.5	33.8	432	2	T06782	extensin precursor
14	235.5	33.8	580	2	T10863	glycine/proline-ri
15	234.5	33.6	173	2	T51469	nodulin precursor
16	233.5	33.5	309	2	S08343	spliceosome-associ
17	233.5	33.5	424	2	A54964	zinc finger protei
18	232.5	33.4	1571	2	T14155	salivary proline-r
19	229.5	32.9	251	1	PIHUPF	proline-rich prote
20	228.5	32.8	188	2	D29149	proline-rich prote
21	225.5	32.4	212	2	B36298	proline-rich prote
22	224	32.1	310	1	PIHUSD	salivary proline-r
23	222.5	31.9	220	2	A36298	proline-rich prote
24	221	31.7	309	2	S10889	proline-rich prote
25	221	31.7	1378	2	T30173	zinc finger protei
26	214.5	30.8	117	2	D40750	proline-rich prote
27	213.5	30.6	301	2	E29149	proline-rich prote
28	210.5	30.2	434	2	S11967	nodule-specific hy
29	210	30.1	297	2	S23737	proline-rich prote

30	210	30.1	440	1	S60755	rhodopsin - Allote
31	210	30.1	620	2	S06733	hydroxyproline-ric
32	209	30.0	97	2	S04554	nodulin - alfalfa
33	208	29.8	223	2	A42817	proline-rich prote
34	207	29.7	727	2	C84534	hypothetical prote
35	206.5	29.6	112	2	S10101	nodulation protein
36	204.5	29.3	539	2	T28770	hypothetical prote
37	202.5	29.1	170	2	A48013	proline-rich prote
38	202	29.0	188	2	JH0481	basic proline-rich
39	201	28.8	452	1	S14332	rhodopsin - northe
40	199.5	28.6	172	2	B29149	proline-rich prote
41	198.5	28.5	1176	2	T49482	hypothetical prote
42	198	28.4	164	2	A30496	proline-rich prote
43	197.5	28.3	260	2	S22373	proline-rich prote
44	197.5	28.3	592	2	T32402	hypothetical prote
45	197.5	28.3	1002	2	T43236	carboxypeptidase C

ALIGNMENTS

RESULT 1

KGHUGH

histidine-rich glycoprotein precursor - human

N;Alternate names: HRG

C;Species: Homo sapiens (man)

C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004

C;Accession: A01287; S296669

R;Koide, T.; Foster, D.; Yoshitake, S.; Davie, E.W.

Biochemistry 25, 2220-2225, 1986

A;Title: Amino acid sequence of human histidine-rich glycoprotein derived from the nucle

A;Reference number: A01287; MUID:86216149; PMID:3011081

A;Accession: A01287

A;Molecule type: mRNA

A;Residues: 1-525 <KOI>

A;Cross-references: UNIPROT:P04196; GB:AB005803; NID:g2280513; PIDN:BAA21613.1; PID:g228

R;Hennis, B.; Havelaar, A.; Klufft, C.

submitted to the EMBL Data Library, October 1991

A;Description: PCR detection of a dinucleotide repeat in the human histidine-rich glycop

A;Reference number: S296669

A;Accession: S296669

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 214-247 <HEN>

A;Cross-references: EMBL:Z17218; NID:g32453; PIDN:CAA78925.1; PID:g32454

C;Comment: Although its physiological function is not yet known, HRG does bind heme, dyes

din, and the lysine-binding site of plasminogen. On the basis of its homology with HMW ki

lood coagulation cascade.

C;Comment: The amino half of this protein is homologous to the first two cystatin-like d

ould not have inhibitory activity.

C;Comment: In addition to having a high histidine and proline content, this protein has n

e-rich' region.

C;Genetics:

A;Gene: GDB:HRG

A;Cross-references: GDB:120055; OMIM:142640

A;Map position: 3q27-3q27

C;Superfamily: histidine-rich glycoprotein; cystatin homology

C;Keywords: duplication; glycoprotein; heparin binding; tandem repeat

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-525/Product: histidine-rich glycoprotein #status predicted <MAT>

F;19-131/Domain: cystatin homology <CY1>

F;140-246/Domain: cystatin homology <CY2>

F;276-321/Region: proline-rich

F;348-437/Region: histidine-rich

F;351-497/Region: proline-rich

F;63,125,344,345/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;78-89,105-126,218-241/Disulfide bonds: #status predicted

Query Match 45.1%; Score 314; DB 1; Length 525;
Best Local Similarity 41.8%; Pred. No. 1.7e-16;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;

QY 7 RPP--PHG-----HHPHGPPPHGHHPHGPPP-----HGHP-PHGPPP----- 40

Db 275 KPPFKPHGSRDHHHPKP-----HEHGPPPPPPDERDHSHPPLPQGPPPLLPMSCSSQH 329

QY 41 -----RHP-----PHGPPPHGHPP-----HGPPPHGHPPHGP 67

Db 330 ATFGTNGAQRHSHNNSSDLPHPKHHSHEQPHGHHPPHAHPHEHDTHRQHPHGHHPHGH 389

QY 68 PPHGHPPPHPPHGHPPHGHGFHDHGPCDPPSH 100

Db 390 HPHGHHPPHGHHPHGHHPHCHDFQDYGYCDPPPH 422

RESULT 2

A60488

histidine-rich glycoprotein - bovine (fragments)

N;Alternate names: autorosette inhibition factor

C;Species: Bos primigenius taurus (cattle)

C;Date: 19-Mar-1993 #sequence_revision 23-Mar-1995 #text_change 07-Jul-1995

C;Accession: S35687; JC2196; A60488

R;Sorensen, C.B.; Krogh-Pedersen, H.; Petersen, T.E.

FEBS Lett. 328, 285-290, 1993

A;Title: Determination of the disulphide bridge arrangement of bovine histidine-rich gly

A;Reference number: S35687; MUID:93351678; PMID:8348977

A;Accession: S35687

A;Molecule type: protein

A;Residues: 1-25;26-52;57-75;82-88;95-119;146-173;175-206;210-309;313-445 <SOR>

A;Note: 355-Gln and 368-Tyr were also found

R;Halkier, T.; Andersen, H.; Vestergaard, A.; Magnusson, S.

Biochem. Biophys. Res. Commun. 200, 78-82, 1994

A;Title: Bovine histidine-rich glycoprotein is a substrate for bovine plasma factor XIII

A;Reference number: JC2196; MUID:94220160; PMID:7909439

A;Accession: JC2196

A;Molecule type: protein

A;Residues: 1-23;35-54,'VK',57-101,'R','TVGEYTEG',116,'N',118,'R',120-136;137-145;150-20

A;Experimental source: plasma

R;Vestergaard, A.B.; Andersen, H.F.; Magnusson, S.; Halkier, T.

Thromb. Res. 60, 385-396, 1990

A;Title: Histidine-rich glycoprotein inhibits contact activation of blood coagulation.

A;Reference number: A60488; MUID:91196010; PMID:2084959

A;Accession: A60488

A;Molecule type: protein

A;Residues: 1-6,'X',8-15 <VES>

C;Comment: This protein is a single-chained plasma protein which participates in transgl

C;Superfamily: histidine-rich glycoprotein; cystatin homology

C;Keywords: glycoprotein; plasma

F;2-113/Domain: cystatin homology <CY1>

F;122-207/Domain: cystatin homology (fragments) <CY2>

F;7-424,60-71,87-108,165-346,180-203,258-288/Disulfide bonds: #status experimental

F;74,107,164,266/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 37.4%; Score 260.5; DB 2; Length 445;

Best Local Similarity 42.4%; Pred. No. 1.2e-12;

Matches 53; Conservative 4; Mismatches 33; Indels 35; Gaps 6;

QY 11 HGHHPHGPPP-----HGHHPHG-----PPPHGHPPHG-----PPP--- 40

Db 229 HEHSPAGRPFFKPSGSKDHG-HPHESYNFRCPPLLEHKNHSDSPFPQARAPLPFFPPGLR 287

QY 41 -RHPHGHPPPHGHPPH--GPPPHGHPPHGHPPHGHPPHGHPPHGHGFHDHGPCD 96

Db 288 CPHPPFGTKGNHRPPPHDSSDEHHNHHPHGHHPHGHHPHGHHPDNDIFYDHGPCD 347

QY 97 PPSHK 101

Db 348 PPHR 352

RESULT 3

B48013

proline-rich proteoglycan 2 precursor, parotid - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 09-Jul-2004

C;Accession: B48013

R;Castle, A.M.; Castle, J.D.

J. Biol. Chem. 268, 20490-20496, 1993

A;Title: Novel secretory proline-rich proteoglycans from rat parotid. Cloning and charact

A;Reference number: A48013; MUID:93388626; PMID:8376404

A;Accession: B48013

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-295 <CAS>

A;Cross-references: UNIPROT:Q07611; GB:L17318; NID:g310199; PIDN:AAA03074.1; PID:g310200

C;Superfamily: proline-rich protein

C;Keywords: extracellular protein; glycoprotein; tandem repeat

Query Match 37.2%; Score 259.5; DB 2; Length 295;

Best Local Similarity 38.4%; Pred. No. 1e-12;

Matches 63; Conservative 1; Mismatches 27; Indels 73; Gaps 10;

QY 2 VNIHRPPPHG---HHPH-----GPPPHG-----HHPHGGPPPHGHP-- 34

Db 81 VNRPERPPQHCGNHHHPHPPPAAGPQRPQPQPGSPQGPFPQGGPQQRPPQGPFPQGGPQR 140

QY 35 -----PHGPPP-----RHPPHGGPPPHGHP-----PHGPPPHGHP-----PHGPPPHG 71

Db 141 PPQPGSPQGPFPFGGPQQRPPQGPFPQGGPQRPQPQPGSPQGPFPFGPQQRAPQGGPPPG 200

QY 72 HP-----PHGP-----PPHGHPPHGHGFHDHGPCDPP 98

Db 201 GPQRPPQPGSPQGPFPFGGPQQRPPQGPFPFG-----GPQRPP 238

RESULT 4

D38355

basic proline-rich peptide IB-8a - human (fragments)

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 09-Jul-2004

C;Accession: D38355

R;Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.

Biochemistry 30, 3351-3356, 1991

A;Title: Basic proline-rich proteins from human parotid saliva: relationships of the cov

A;Reference number: A38355; MUID:91190884; PMID:1849422

A;Accession: D38355

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-128 <KAU>

A;Cross-references: UNIPROT:Q7M4Q5

C;Superfamily: proline-rich protein

Query Match 35.9%; Score 250; DB 2; Length 128;

Best Local Similarity 45.0%; Pred. No. 2.5e-12;

Matches 50; Conservative 5; Mismatches 36; Indels 20; Gaps 5;

QY 8 PPHGHHPHGGPPHGHHPHGHPPH-GHPPHGGPPPRHPPHGHPPPHGH-----PPHGGPPPH 60

Db 10 PPQGNQPPQGPFPFPFGPKPQGPFPQGGNKPQGPFPFPFGPKPQGPFPQGDNKSQSAKPQGPFPQ 69

QY 61 GHPPHGGPPPH-GHPPHGGPPPHGH-----PPHGHGFHDHGPCDPPS 99

Db 70 GKKPQGPFPQCGNKPQGPFPQGSARSFPCKPQGPFPQEQEGNNPQGP-PPPA 119

RESULT 5

B39066

proline-rich protein 15 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Mar-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004

C;Accession: B39066

R;Lin, H.H.; Ann, D.K.

Genomics 10, 102-113, 1991

A;Title: Molecular characterization of rat multigene family encoding proline-rich protei

A;Reference number: A39066; MUID:91257817; PMID:2045095

A;Accession: B39066

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-278 <LIN>

QY	7	RPPPHGHHP-----HGPPPHG-----HHPHGPPPHG-----HPPHGPPP-----RHPPH	45
Db	33	RPPPSGSQRP RPVNGSQGGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQ	92
QY	46	GPPPHG-----HPPHGPPPHG-----HPPHGPPPHG-----HPPHGPPPHG-----HPPHG	86
Db	93	GPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPP	152
QY	87	-HGFHDHGPCDPP	98
Db	153	PGGPQRP RP PQGPP	165
RESULT 10			
B24264			
proline-rich protein MP3 - mouse (fragment)			
C;Species: Mus musculus (house mouse)			
C;Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 03-May-1996			
C;Accession: B24264			
R;Ann, D.K.; Carlson, D.M.			
J. Biol. Chem. 260, 15863-15872, 1985			
A;Title: The structure and organization of a proline-rich protein gene of a mouse multi			
A;Reference number: A92508; MUID:86059475; PMID:2999141			
A;Accession: B24264			
A;Molecule type: DNA			
A;Residues: 1-240 <ANN>			
C;Superfamily: proline-rich protein			
Query Match 34.3%; Score 239; DB 2; Length 240;			
Best Local Similarity 42.6%; Pred. No. 2.7e-11;			
Matches 58; Conservative 0; Mismatches 32; Indels 46; Gaps 10;			
QY	7	RPPPHGHHP-----HGPPPHG-----HHPHGPPPHG-----HPPHGPPP-----RHPPH	45
Db	12	RPPPSGSQRP RPVNGSQGGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQ	71
QY	46	GPPPHG-----HPPHGPPPHG-----HPPHGPPPHG-----HPPHGPP-----PHGHPPHG---	86
Db	72	GPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPPPPGGPQRP RP PQGPPPPAGPQRP RP PQGPP	131
QY	87	----HGFHDHGPCDPP	98
Db	132	PAGPHLRPTQGP--PP	145
RESULT 11			
S22140			
nodulin Enod2 - Sesbania rostrata			
C;Species: Sesbania rostrata			
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004			
C;Accession: S22140			
R;de Bruijn, F.J.			
submitted to the EMBL Data Library, November 1991			
A;Reference number: S22140			
A;Accession: S22140			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-330 <DEB>			
A;Cross-references: UNIPROT:Q41402; EMBL:X63339; NID:g21372; PIDN:CAA44939.1; PID:g21373			
C;Superfamily: proline-rich protein 3			
Query Match 33.9%; Score 236; DB 2; Length 330;			
Best Local Similarity 41.1%; Pred. No. 5.8e-11;			
Matches 44; Conservative 10; Mismatches 41; Indels 12; Gaps 3;			
QY	7	RPPPHGHHPGPPPHGHHPHG-----PPPHGHPPHGPppRH---PPHGPPPHGHPPHG	56
		:	
Db	151	KPPEYQPPHEKPPPEYQPPQEQKPPPVYPPPYEKPHEKPPPSYEKPPPYEKPHEKPPPYE	210
QY	57	PPPHGHPPHGPppPHGHPPHGPppPHGHPPHGHGFDHGPCD--PPSHK	101
Db	211	KPHEKPPHEKPPYDKPPPYEKPHEKPPHEKPPHEKPPPYEKPHEK	257

RESULT 12

PIHUB6

salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - human
N;Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: B40750; C40750; A40750; C25372; S02128; S02127; A03293; A90502; A91974; A052
R;Azen, E.A.; Latreille, P.; Niece, R.L.
Am. J. Hum. Genet. 53, 264-278, 1993
A;Title: PRB1 gene variants coding for length and null polymorphisms among human salivary
A;Reference number: A40750; MUID:93304421; PMID:8317492
A;Accession: B40750
A;Molecule type: DNA
A;Residues: 35-392 <AZE>
A;Cross-references: UNIPROT:P04280; UNIPROT:P04281; UNIPROT:P02811; UNIPROT:Q16038; GB:S
A;Experimental source: subject C.J. (large allele)
A;Accession: C40750
A;Molecule type: DNA
A;Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-392 <
A;Cross-references: GB:S62929
A;Experimental source: subject M.V.O. (large allele)
A;Accession: A40750
A;Molecule type: DNA
A;Residues: 35-183,245-270,'Q',272-392 <AZ3>
A;Cross-references: GB:S62928
A;Experimental source: subject C.J. (medium allele)
A;Note: authors translated the codon CAA for residue 272 as Arg
R;Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985
A;Title: Differential RNA splicing and post-translational cleavages in the human salivary
A;Reference number: A92492; MUID:85289325; PMID:2993301
A;Accession: C25372
A;Molecule type: mRNA
A;Residues: 1-183,245-392 <MAE>
A;Cross-references: GB:X03204; NID:g190485; PIDN:AAA60185.1; PID:g190486
A;Note: alternatively splice forms lacking portions of the repeat region were also found
R;Lyons, K.M.; Stein, J.H.; Smithies, O.
Genetics 120, 267-278, 1988
A;Title: Length polymorphisms in human proline-rich protein genes generated by intragenic
A;Reference number: S02127; MUID:89121440; PMID:2851479
A;Accession: S02128
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>
A;Cross-references: EMBL:X07517
A;Accession: S02127
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 35-183,245-392 <LY2>
A;Cross-references: EMBL:X07516
R;Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P.
Biochemistry 25, 2387-2392, 1986
A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent structu
A;Reference number: A90502; MUID:86243355; PMID:3521730
A;Accession: A03293
A;Molecule type: protein
A;Residues: 17-38,'AP',41-51,92-148,'R',150-152 <KA2>
A;Note: among nine basic proline-rich peptides isolated from the saliva, this peptide is
A;Accession: A90502
A;Molecule type: protein
A;Residues: 275-336,'S',338-392 <KAU>
R;Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 94, 1991-1999, 1983
A;Title: Further fractionation of basic proline-rich peptides from human parotid saliva
A;Reference number: A91974; MUID:84161824; PMID:6671974
A;Contents: P-H
A;Accession: A91974
A;Molecule type: protein
A;Residues: 'S',338-392 <SAI>
R;Azen, E.; Lyons, K.M.; McGonigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Vanin,
Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984

A;Reference number: A94005; MUID:84298176; PMID:6089212
A;Accession: A05261
A;Molecule type: DNA
A;Residues: 35-39,'P',41-84,'G',86,'R',87-154,'R',218-246;300-306,'T',308-329,'C',331-384
A;Accession: A05262
A;Molecule type: DNA
A;Residues: 'N',57-59,'A',61-69;334-336,'S',338-339,'R',341-392 <AZ5>
R;Kauffman, D.; Wong, R.; Bennick, A.; Keller, P.
Biochemistry 21, 6558-6562, 1982
A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent structure
A;Reference number: A90464; MUID:83101329; PMID:5924859
A;Contents: IB-9
A;Accession: A90464
A;Molecule type: protein
A;Residues: 92-127,'R',129-148,'R',150-152 <KA3>
R;Isemura, S.; Saitoh, E.; Sanada, K.
J. Biochem. 91, 2067-2075, 1982
A;Title: Fractionation and characterization of basic proline-rich peptides of human parotid
A;Reference number: A91966; MUID:83007119; PMID:7118863
A;Contents: P-E
A;Accession: A91966
A;Molecule type: protein
A;Residues: 92-127,'R',129-148,'R',150-152 <ISE>
C;Comment: This peptide contains 21-residue repeats, two of which have internal 7-residue
C;Genetics:
A;Gene: GDB:PRB1
A;Cross-references: GDB:119511; OMIM:180989
A;Map position: 12p13.2-12p13.2
A;Note: each of the tandem repeats contains a candidate splice acceptor site, and several
C;Superfamily: proline-rich protein
C;Keywords: alternative splicing; duplication; parotid gland; phosphoprotein; pyroglutamate
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-51,92-152/Product: basic proline-rich peptide IB-1 #status experimental <IB1>
F;92-152/Product: basic proline-rich peptide P-E #status experimental <PPE>
F;275-392/Product: basic proline-rich peptide IB-6 #status experimental <PIB6>
F;275-335/Product: basic proline-rich peptide P-F #status experimental <PPF>
F;337-392/Product: basic proline-rich peptide P-H #status experimental <PPH>
F;17/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
F;24/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 33.8%; Score 235.5; DB 1; Length 392;
Best Local Similarity 41.4%; Pred. No. 7.3e-11;
Matches 48; Conservative 4; Mismatches 39; Indels 25; Gaps 4;

Qy 8 PPHGHHPGPPPHGHHPHGGPPPH-GHPPHGPPRHRPPHGGPPH----- 50
Db 101 PPOGGNQPGPPPPPGKPGQPPPPQGGNKPGQPPPPGKPGQPPPPQGDKSQSPRSPGKPGQ 160

Qy 51 -----GHPPHGGPPPHGPPHGGPPPH-GHPPHGPPPHGHHP--PHGHGFHDHGPCDPP 98
Db 161 PPOGGNQPGPPPPPGKPGQPPPPQGGNKPGQPPPPGKPGQPPPPQGDKSQSPRSP 216

RESULT 13
T06782
extensin - soybean
N;Alternate names: hydroxyproline-rich glycoprotein
C;Species: Glycine max (soybean)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06782
R;Ahn, J.H.; Choi, Y.; Kwon, Y.M.; Kim, S.G.; Choi, Y.D.; Lee, J.S.
Plant Cell 8, 1477-1490, 1996
A;Title: A novel extensin gene encoding a hydroxyproline-rich glycoprotein requires sucrose
A;Reference number: Z15809; MUID:96434536; PMID:8837503
A;Accession: T06782
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-432 <AHN>
A;Cross-references: UNIPROT:Q39835; EMBL:U44838; NID:g1165321; PIDN:AAB53156.1; PID:g1165321
C;Genetics:
A;Gene: HRGP3
C;Superfamily: proline-rich protein 3
C;Keywords: glycoprotein; hydroxyproline

Query Match 33.8%; Score 235.5; DB 2; Length 432;
Best Local Similarity 41.7%; Pred. No. 8e-11;
Matches 58; Conservative 2; Mismatches 32; Indels 47; Gaps 14;

Qy 6 HRPPPHGHHP-----HGPPPHGHHP-----HGPPP--HGHP-----HGPPRHRPPHGP 47
Db 49 HSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPP--PKHSP 106

Qy 48 PP-----HGHP--HGPPP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP 85
Db 107 PPPYYHSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPP 166

Qy 86 GHG----FHDHGPCDPPSH 100
Db 167 KHSPPPPYYHSP-PPPKH 184

RESULT 14
T10863
extensin precursor - kidney bean
N;Alternate names: cell wall protein; hydroxyproline-rich glycoprotein
C;Species: Phaseolus vulgaris (kidney bean)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10863; B29356
R;Wycoff, K.L.; Powell, P.A.; Gonzales, R.A.; Corbin, D.R.; Lamb, C.; Dixon, R.A.
Plant Physiol. 109, 41-52, 1995
A;Title: Stress activation of a bean hydroxyproline-rich glycoprotein promoter is superin
A;Reference number: Z17192; MUID:96061709; PMID:7480331
A;Accession: T10863
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-580 <WYC>
A;Cross-references: UNIPROT:Q09083; EMBL:U18791; NID:g727263; PIDN:AAA87902.1; PID:g727263
A;Experimental source: strain Tendergreen
R;Corbin, D.R.; Sauer, N.; Lamb, C.J.
Mol. Cell. Biol. 7, 4337-4344, 1987
A;Title: Differential regulation of a hydroxyproline-rich glycoprotein gene family in wo
A;Reference number: A29356; MUID:88142825; PMID:3437892
A;Accession: B29356
A;Molecule type: mRNA
A;Residues: 'N',1-229 <COR>
A;Cross-references: EMBL:M18094; NID:g169346; PIDN:AAA33764.1; PID:g169347
A;Experimental source: cv. Kievitsboon Koskoek
C;Superfamily: hydroxyproline-rich glycoprotein
C;Keywords: glycoprotein; hydroxyproline
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-580/Product: extensin #status predicted <MAT>

Query Match 33.8%; Score 235.5; DB 2; Length 580;
Best Local Similarity 41.7%; Pred. No. 1e-10;
Matches 58; Conservative 2; Mismatches 32; Indels 47; Gaps 14;

Qy 6 HRPPPHGHHP-----HGPPPHGHHP-----HGPPP--HGHP-----HGPPRHRPPHGP 47
Db 61 HSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPP--PKHSP 118

Qy 48 PP-----HGHP--HGPPP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP 85
Db 119 PPPYYHSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPPPKHSPPPPPPYYHSPPP 178

Qy 86 GHG----FHDHGPCDPPSH 100
Db 179 KHSPPPPYYHSP-PPPKH 196

RESULT 15
T51469
glycine/proline-rich protein - Arabidopsis thaliana
N;Alternate names: protein K10A8_130
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51469

R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51469
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <SAT>
A;Cross-references: UNIPROT:Q9LF59; EMBL:AL391151
A;Experimental source: cultivar Columbia; BAC clone K10A8
C;Genetics:
A;Map position: 5
A;Introns: 97/1
A;Note: K10A8_130

Query Match	33.6%	Score	234.5;	DB 2;	Length	173;			
Best Local Similarity	51.7%	Pred. No.	4.3e-11;						
Matches	45;	Conservative	7;	Mismatches	22;	Indels	13;	Gaps	5;

Qy	9	PPHGHHPHGPPPHGHHPHGPPPHGHP-PHGGPPRRPPHGGPPPHGHPPHGP-PPHGHPPHG	66
		: : : : : : :	
Db	26	PPHGH-----GYGHHGHG---YGSSYPYPPPP--PPHGYPPVAYPPHGGYPYPAGYPPAG	74

Qy	67	PPPHGHPPHGGPPPHGHPPHGHGHFDHG	93
		: : : : : :	
Db	75	YPPAGYPAHGYPSHGYPRPSHSGHHG	101

Search completed: June 15, 2005, 13:51:56
Job time : 20.4699 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:32:29 ; Search time 90.8594 Seconds
(without alignments)
569.231 Million cell updates/sec

Title: US-10-074-225A-6
Perfect score: 697
Sequence: 1 SVNIIHRPPPHGHGHPGPP.....HPPHGHGFHDHGPCDPPSHK 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	697	100.0	526	1	HRG_RABIT	Q28640 oryctolagus
2	420.5	60.3	546	2	Q9BGU1	Q9bgul bos taurus
3	375	53.8	525	2	Q99PS8	Q99ps8 rattus norv
4	333	47.8	525	2	Q99PS5	Q99ps5 mus musculu
5	333	47.8	525	2	Q99PS6	Q99ps6 mus musculu
6	333	47.8	525	2	Q9ESB3	Q9esb3 mus musculu
7	333	47.8	536	2	Q6YK32	Q6yk32 mus musculu
8	333	47.8	536	2	Q6YKA2	Q6yka2 mus musculu
9	323	46.3	515	2	Q99PS7	Q99ps7 rattus norv
10	315	45.2	510	2	Q9ESB2	Q9esb2 rattus norv
11	314	45.1	312	2	Q68DR3	Q68dr3 homo sapien
12	314	45.1	525	1	HRG_HUMAN	P04196 homo sapien
13	260	37.3	396	1	HRG_BOVIN	P33433 bos taurus
14	259.5	37.2	295	2	Q07611	Q07611 rattus norv
15	258.5	37.1	511	2	Q95JD0	Q95jd0 sus scrofa
16	258.5	37.1	566	2	Q95JD1	Q95jd1 sus scrofa
17	254.5	36.5	676	2	Q95JC9	Q95jc9 sus scrofa
18	250	35.9	128	2	Q7M4Q5	Q7m4q5 homo sapien
19	249	35.7	1109	2	Q7SFA3	Q7sfa3 neurospora
20	247.5	35.5	206	1	PRP3_RAT	P04474 rattus norv
21	246.5	35.4	227	2	Q62107	Q62107 mus musculu
22	246.5	35.4	261	1	PRP2_MOUSE	P05142 mus musculu
23	246.5	35.4	296	1	PMP3_MOUSE	P05143 mus musculu
24	246.5	35.4	317	2	Q62103	Q62103 mus musculu
25	246	35.3	274	2	Q04154	Q04154 rattus norv
26	240	34.4	1646	2	Q7PRH5	Q7prh5 anopheles g
27	236	33.9	257	2	Q615I7	Q615i7 oryza sativ
28	236	33.9	330	2	Q41402	Q41402 sesbania ro
29	235.5	33.8	230	2	Q7DMV8	Q7dmv8 phaseolus v
30	235.5	33.8	392	1	PRP1_HUMAN	P04280 homo sapien
31	235.5	33.8	432	2	Q39835	Q39835 glycine max

32	235.5	33.8	580	2	Q09083	Q09083 phaseolus v
33	234.5	33.6	173	2	Q9LF59	Q9lf59 arabidopsis
34	233.5	33.5	309	1	NO75_SOYEN	P08297 glycine max
35	233.5	33.5	338	2	Q86YA1	Q86ya1 homo sapien
36	233.5	33.5	424	1	S3B4_HUMAN	Q15427 homo sapien
37	233.5	33.5	424	2	Q8QZY9	Q8qzy9 m splicing
38	233.5	33.5	424	2	Q6AYL5	Q6ayl5 rattus norv
39	233.5	33.5	586	2	Q6CAY8	Q6cay8 yarrowia li
40	232.5	33.4	1571	2	O54978	O54978 mus musculu
41	232.5	33.4	1572	2	Q8RSN0	Q8r5n0 mus musculu
42	232.5	33.4	1617	2	Q80U47	Q80u47 mus musculu
43	231	33.1	173	2	Q15214	Q15214 homo sapien
44	230.5	33.1	382	1	PRB2_HUMAN	P02812 homo sapien
45	230	33.0	494	2	Q68D96	Q68d96 homo sapien

ALIGNMENTS

RESULT 1

HRG_RABIT	STANDARD;	PRT;	526 AA.
AC Q28640;			
DT 01-NOV-1997 (Rel. 35, Created)			
DT 01-NOV-1997 (Rel. 35, Last sequence update)			
DT 05-JUL-2004 (Rel. 44, Last annotation update)			
DE Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG) (Fragment).			
GN Name=HRG;			
OS Oryctolagus cuniculus (Rabbit).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX NCBI_TaxID=9986;			
RN [1]			
RP SEQUENCE FROM N.A., AND SEQUENCE OF 9-23; 301-313 AND 422-429.			
RC TISSUE=Serum;			
RX MEDLINE=96229917; PubMed=8639676; DOI=10.1021/bi952061t;			
RA Borza D.-B., Tatum F.M., Morgan W.T.;			
RT "Domain structure and conformation of histidine-proline-rich glycoprotein.";			
RL Biochemistry 35:1925-1934(1996).			
CC -!- FUNCTION: The physiological function is not yet known. It binds heme, dyes and divalent metal ions. It can inhibit rosette formation and is known to interact with heparin, thrombospondin, and the lysine-binding site of plasminogen. On the basis of its homology with HMW kininogen, the His-rich region of this protein may mediate the contact activation phase of intrinsic blood coagulation cascade.			
CC -!- SUBCELLULAR LOCATION: Secreted.			
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.			
CC -!- DOMAIN: In addition to having a high His and Pro content, this protein has many internal repeats. 15 tandem repetitions of a 5-residue sequence (GH/PPH, consensus) form a His/Pro-rich region.			
CC -!- SIMILARITY: Contains 2 cystatin-like domains.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL; U32189; AAC48516.1; --			
DR InterPro; IPR000010; Prot_inh_cystat.			
DR Pfam; PF00031; Cystatin; 1.			
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat; Signal.			
FT NON_TER	1	1	Potential.
FT SIGNAL	<1	8	
FT CHAIN	9	526	Histidine-rich glycoprotein.
FT DOMAIN	9	126	Cystatin-like 1.

FT	DOMAIN	127	243	Cystatin-like 2.
FT	DOMAIN	251	296	Pro-rich.
FT	DOMAIN	329	498	His/Pro-rich.
FT	DISULFID	14	505	By similarity.
FT	DISULFID	68	79	By similarity.
FT	DISULFID	95	116	By similarity.
FT	DISULFID	193	415	By similarity.
FT	DISULFID	207	230	By similarity.
FT	DISULFID	272	302	Potential.
FT	CARBOHYD	115	115	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	192	192	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	240	240	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	310	310	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	485	485	N-linked (GlcNAc. . .) (Potential).
FT	SITE	303	304	Cleavage (by plasmin).
FT	SITE	421	422	Cleavage (by plasmin).
SQ	SEQUENCE	526 AA;	58877 MW;	810F23D367D93D42 CRC64;
Query Match 100.0%; Score 697; DB 1; Length 526;				
Best Local Similarity 100.0%; Pred. No. 2.6e-38;				
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	SVNIHRPPPHGHPHGGPPPHGHHPHGGPPPHGHHPHGGPPPPHGGPPPHGGPPPH 60		
Db	321	SVNIHRPPPHGHPHGGPPPHGHHPHGGPPPHGHHPHGGPPPPHGGPPPHGGPPPH 380		
Qy	61	GHPHGGPPPHGHPPPHGHPHGGPPPHGHGHGFDHGGPCDPPSHK 101		
Db	381	GHPHGGPPPHGHPHGGPPPHGHHPHGGPPPHGHGHGFDHGGPCDPPSHK 421		
RESULT 2				
ID	Q9BGU1	PRELIMINARY;	PRT;	546 AA.
AC	Q9BGU1;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	Histidine-rich glycoprotein.			
GN	Name=BTHRG;			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,			
RA	Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB055894; BAB33091.1; --			
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR00010; Prot_inh_cystat.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; CY; 2.			
SQ	SEQUENCE	546 AA;	61948 MW;	26264858824D89EE CRC64;
Query Match 60.1%; Score 420.5; DB 2; Length 546;				
Best Local Similarity 63.8%; Pred. No. 2.6e-20;				
Matches 67; Conservative 4; Mismatches 21; Indels 13; Gaps 2;				
Qy	6	HRPP-----PHGHHPPPHGHPHGGPPPHGHPHGGPPPPHGGPPPHGGPPPHG 56		
Db	341	HRPPHDHSSDEHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 396		
Qy	57	PPPHGHPHGGPPPHGHPHGGPPPHGHPHGGPPPHGHPHGGPPPPHGGPPSHK 101		
Db	397	HHPHGHHPHGHHPHGHHPYGHHPHGHHPHHPNDYFDHGGPCDPPPHR 441		
RESULT 3				
ID	Q99PS8	PRELIMINARY;	PRT;	525 AA.

AC	Q99PS8;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	Histidine-rich glycoprotein 1.			
GN	Name=RNHRG1;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Liver;			
RA	Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,			
RA	Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB055895; BAB33092.1; --			
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.			
DR	InterPro; IPR00010; Prot_inh_cystat.			
DR	Pfam; PF00031; Cystatin; 1.			
DR	SMART; SM00043; CY; 2.			
SQ	SEQUENCE	525 AA;	59049 MW;	38290A631FAC7777 CRC64;
Query Match 53.8%; Score 375; DB 2; Length 525;				
Best Local Similarity 63.5%; Pred. No. 2.3e-17;				
Matches 66; Conservative 2; Mismatches 24; Indels 12; Gaps 5;				
Qy	3	NIHRPPPHGH---HP-HGPPPHGHHPHGGPPPHGHHPHGGPPR-HPPHGGPPHGHPPHG 56		
Db	322	NRSHR-PPNHSCNEHPCHGQHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 380		
Qy	57	PPPHGHPPHGGPPPHGHPHGGPPPHGHHPHGHGHGFDHGGPCDPPSH 100		
Db	381	QHPHGH-----HPPHGHHPHGHHPHGHHPHGHDFLDYGPCDPPSN 419		
RESULT 4				
ID	Q99PS5	PRELIMINARY;	PRT;	525 AA.
AC	Q99PS5;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)		
DE	Histidine-rich glycoprotein.			
GN	Name=Hrg; Synonyms=MMHRG;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tsuchida N., Wakabayashi S., Jahnen-Dechent W., Koide T.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N; TISSUE=Liver;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fabey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			

```

RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055898; BAB33095.1; -.
DR EMBL; BC011168; AAH11168.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59162 MW; A83E93A439CFB126 CRC64;

Query Match 47.8%; Score 333; DB 2; Length 525;
Best Local Similarity 58.3%; Pred. No. 1.2e-14;
Matches 60; Conservative 3; Mismatches 30; Indels 10; Gaps 4;

Qy 3 NIIHRPPPHGH---HP-HGPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 57
Db 322 NRTHR-PSYNHSCNEHPCHGHRPHGHHPHSHHPGHHSHG---HHPHGHHPHSHHSHGH 376

Qy 58 PPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 100
Db 377 HPPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 419

RESULT 5
Q99PS6 PRELIMINARY; PRT; 525 AA.
AC Q99PS6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg; Synonyms=MMHRG;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055897; BAB33094.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59090 MW; A83E93A439CFB3AC CRC64;

Query Match 47.8%; Score 333; DB 2; Length 525;
Best Local Similarity 58.3%; Pred. No. 1.2e-14;
Matches 60; Conservative 3; Mismatches 30; Indels 10; Gaps 4;

Qy 3 NIIHRPPPHGH---HP-HGPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 57
Db 322 NRTHR-PSYNHSCNEHPCHGHRPHGHHPHSHHPGHHSHG---HHPHGHHPHSHHSHGH 376

Qy 58 PPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 100
Db 377 HPPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 419

RESULT 6
Q9ESB3 PRELIMINARY; PRT; 525 AA.
AC Q9ESB3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=20307726; PubMed=10849117;
RA Hulett M.D., Parish C.R.;
RT "Murine histidine-rich glycoprotein: cloning, characterization and
RT cellular origin.";
RL Immunol. Cell Biol. 78:280-287(2000).
DR EMBL; AF194028; AAG28416.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59132 MW; 6E55F2A439CFB123 CRC64;

Query Match 47.8%; Score 333; DB 2; Length 525;
Best Local Similarity 58.3%; Pred. No. 1.2e-14;
Matches 60; Conservative 3; Mismatches 30; Indels 10; Gaps 4;

Qy 3 NIIHRPPPHGH---HP-HGPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 57
Db 322 NRTHR-PSYNHSCNEHPCHGHRPHGHHPHSHHPGHHSHG---HHPHGHHPHSHHSHGH 376

Qy 58 PPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 100
Db 377 HPPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 419

RESULT 7
Q6YK32 PRELIMINARY; PRT; 536 AA.
AC Q6YK32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Histidine-rich glycoprotein HRG.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137504; AAN27996.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR00010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 536 AA; 60439 MW; CCF14FC08DD53D7C CRC64;

Query Match 47.8%; Score 333; DB 2; Length 536;
Best Local Similarity 58.3%; Pred. No. 1.3e-14;
Matches 60; Conservative 3; Mismatches 30; Indels 10; Gaps 4;

Qy 3 NIIHRPPPHGH---HP-HGPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 57
Db 333 NRTHR-PSYNHSCNEHPCHGHRPHGHHPHSHHPGHHSHG---HHPHGHHPHSHHSHGH 387

Qy 58 PPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGHPPPHGP 100
```


Job time : 92.8594 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-7

Perfect score: 23

Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	91.3	4	5	ABG77500	Abg77500 Targetin
2	21	91.3	5	2	AAR12701	Aar12701 Pentapept
3	21	91.3	5	2	AAR51565	Aar51565 Mimotope
4	21	91.3	5	2	AAR69933	Aar69933 Pentameri
5	21	91.3	5	2	AAR89833	Aar89833 Melanotro
6	21	91.3	5	2	AAR98679	Aar98679 Peptide 5
7	21	91.3	5	2	AAW78934	Aaw78934 Antidepre
8	21	91.3	5	2	AAY23419	Aay23419 V beta 6
9	21	91.3	5	5	ABB79810	Abb79810 Histidine
10	21	91.3	5	5	ABB79808	Abb79808 Histidine
11	21	91.3	5	5	ABB79809	Abb79809 Histidine
12	21	91.3	5	5	ABB79811	Abb79811 Histidine
13	21	91.3	5	5	AAG80338	Aag80338 Enzyme cl
14	21	91.3	5	5	AAO15016	Aao15016 Mutant VL
15	21	91.3	5	8	ADH10419	Adh10419 HPRG prot
16	21	91.3	5	8	ADH10422	Adh10422 HPRG prot
17	21	91.3	5	8	ADH10417	Adh10417 Rabbit HP
18	21	91.3	5	8	ADH10415	Adh10415 Rabbit HP
19	21	91.3	5	8	ADH10416	Adh10416 Rabbit HP
20	21	91.3	5	8	ADH10418	Adh10418 HPRG prot
21	21	91.3	5	8	ADO81122	Ado81122 Sheep pri
22	21	91.3	5	8	ADR03627	Adr03627 E. coli m
23	21	91.3	6	2	AAW75370	Aaw75370 Hexapepti
24	21	91.3	6	2	AAW75302	Aaw75302 Hexapepti
25	21	91.3	6	4	AAB55516	Aab55516 Human ela

26	21	91.3	6	5	ABG31535	Abg31535 Camptothe
27	21	91.3	6	5	AAE26391	Aae26391 Human GPR
28	21	91.3	6	5	ABP53526	Abp53526 Camptothe
29	21	91.3	6	7	ADN12399	Adn12399 Peptide #
30	21	91.3	6	8	ADO37172	Ado37172 Binding p
31	21	91.3	6	8	ADO37904	Ado37904 Binding p
32	21	91.3	6	8	ADO37705	Ado37705 Binding p
33	21	91.3	6	8	ADO37180	Ado37180 Binding p
34	21	91.3	6	8	ADO37698	Ado37698 Binding p
35	21	91.3	6	8	ADO37897	Ado37897 Binding p
36	21	91.3	6	8	ADO28227	Ado28227 Capture s
37	21	91.3	6	8	ADO28419	Ado28419 Capture s
38	21	91.3	6	8	ADO28426	Ado28426 Capture s
39	21	91.3	6	8	ADO27694	Ado27694 Capture s
40	21	91.3	6	8	ADO27702	Ado27702 Capture s
41	21	91.3	6	8	ADO28220	Ado28220 Capture s
42	21	91.3	6	8	ADR49685	Adr49685 6-mer pep
43	21	91.3	6	8	ADR50404	Adr50404 6-mer pep
44	21	91.3	6	8	ADR49879	Adr49879 6-mer pep
45	21	91.3	6	8	ADR49871	Adr49871 6-mer pep

ALIGNMENTS

RESULT 1
ABG77500
ID ABG77500 standard; peptide; 4 AA.
XX
AC ABG77500;
XX

DT 05-NOV-2002 (first entry)
XX

DE Targetting peptide selective for human organ, tissue or cell type #33.
XX

KW Human; cytostatic; antiinflammatory; antidiabetic; cardiovascular;
KW immunomodulator; antibacterial; antiviral; gene therapy; cancer;
KW arthritis; diabetes; inflammatory disease; atherosclerosis;
KW autoimmune disease; bacterial infection; viral infection;
KW cardiovascular disease; degenerative disease.
XX

OS Homo sapiens.
XX

PN WO200220723-A2.
XX

PD 14-MAR-2002.
XX

PF 07-SEP-2001; 2001WO-US028044.
XX

PR 08-SEP-2000; 2000US-0231266P.
PR 17-JAN-2001; 2001US-00765101.

PA (TEXA) UNIV TEXAS SYSTEM.
XX

PI Arap W, Pasqualini R;
XX

DR WPI; 2002-599247/64.
XX

PT New targeting peptides identified by phage display, useful for treating a
PT disease state, e.g. cancer, diabetes, inflammatory disease,
PT atherosclerosis, autoimmune disease, bacterial or viral infection or
PT cardiovascular disease.
XX

PS Claim 16; Page 63; 269pp; English.
XX

CC The invention describes an isolated peptide of 100 amino acids or less in
CC size. The peptide is useful for treating a disease state, e.g. cancer,
CC arthritis, diabetes, inflammatory disease, atherosclerosis, autoimmune
CC disease, bacterial infection, viral infection, cardiovascular disease or
CC degenerative disease. This sequence represents a human targeting peptide
XX selective for human organs, tissues or cell types

SQ Sequence 4 AA;

Query Match									
Best Local Similarity 91.3%; Score 21; DB 5; Length 4;									
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	3	PHG 5							
Db	1	PHG 3							
RESULT 2									
AAR12701									
ID	AAR12701 standard; protein; 5 AA.								
XX									
AC	AAR12701;								
XX									
DT	31-JUL-1991 (first entry)								
XX									
DE	Pentapeptide paralogue for use as chromatographic affinity ligand.								
XX									
KW	HPLC.								
XX									
OS	Synthetic.								
XX									
PN	WO9106356-A.								
XX									
PD	16-MAY-1991.								
XX									
PF	31-OCT-1989; 89US-00429721.								
XX									
PR	31-OCT-1989; 89US-00429721.								
XX									
PA	(TERR-) TERRAPIN TECHN INC.								
XX									
PI	Kauvar LM;								
XX									
DR	WPI; 1991-163985/22.								
XX									
PT	Identifying paralogueueueue with specific affinity for analyte - using								
PT	candidate paralogues with systematically varied values of at least 2								
PT	parameters.								
XX									
PS	Claim 41; Fig 8; 95pp; English.								
XX									
CC	Peptide is one of a panel, each of which have systemically varied values of								
CC	at least two parameters. A test protein is matched against the panel, and								
CC	the paralogues with highest binding affinity are selected. Selected								
CC	paralogues are bound to a matrix which is incorporated into the HPLC gel								
CC	and used as affinity ligands. Technique is useful in chromatographic								
CC	separation, purification and binding assay								
XX									
SQ	Sequence 5 AA;								
Query Match 91.3%; Score 21; DB 2; Length 5;									
Best Local Similarity 100.0%; Pred. No. 1.8e+06;									
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	3	PHG 5							
Db	2	PHG 4							
RESULT 3									
AARS1565									
ID	AARS1565 standard; peptide; 5 AA.								
XX									
AC	AARS1565;								
XX									
DT	25-MAR-2003 (revised)								
DT	19-OCT-1994 (first entry)								
XX									
DE	Mimotope peptide #59 from panel of maximally diverse mimotopes.								
XX									
Query Match 91.3%; Score 21; DB 2; Length 5;									
Best Local Similarity 100.0%; Pred. No. 1.8e+06;									
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	3	PHG 5							
Db	2	PHG 4							
RESULT 4									
AAR69933									
ID	AAR69933 standard; peptide; 5 AA.								
XX									
AC	AAR69933;								
XX									
DT	25-MAR-2003 (revised)								
DT	18-OCT-1995 (first entry)								
XX									
DE	Pentameric mimotope 59 used to obtain highly specific antibodies.								
XX									
KW	mimotope; antibody; production; high specificity; detection; immunoassay;								
KW	high performance liquid chromatography.								
XX									
OS	Synthetic.								
XX									
FH	Key Location/Qualifiers								
FT	Modified-site 1								
FT	/note= "Acetyl-His"								
mimotope panel; rational drug design; candidate drug; screening assay;									
hydrophobicity; antibody repertoire.									
XX									
OS	Synthetic.								
XX									
FH	Key Location/Qualifiers								
FT	Modified-site 1								
FT	/note= "acetylated"								
XX									
PN	US5300425-A.								
XX									
PD	05-APR-1994.								
XX									
PF	06-DEC-1989; 89US-00447009.								
XX									
PR	13-OCT-1987; 87US-00108130.								
PR	11-OCT-1988; 88US-00255906.								
XX									
PA	(TERR-) TERRAPIN TECHNOLOGIES INC.								
XX									
PI	Kauvar LM;								
XX									
DR	WPI; 1994-109390/13.								
XX									
PT	Screening of candidate drugs for binding to receptor - by comparing								
PT	inverse image antibody profile of drug with mimotope-binding profile of								
PT	receptor.								
XX									
PS	Example 3; Fig 3; 29pp; English.								
XX									
CC	A panel of 88 pentapeptides was designed on the basis of decreasing								
CC	hydrophobicity and periodic variation of hydrophobic moment. The peptides								
CC	were labelled with iodine-125 and tested with individual members of a								
CC	basal antibody repertoire. (Spleen cells were harvested from mice and								
CC	used to provide a panel of antibody-secreting hybridoma cells as a subset								
CC	of the complete B-cell repertoire). Nearly uniform binding to all								
CC	antibody members of the repertoire was observed. The test was than								
CC	repeated with the addition of a defined amount of analyte to the mixture.								
CC	A small number of wells showed greatly decreased labelling and these								
CC	antibodies represented the successful result of an initial screen for								
CC	those which preferentially bind analyte. (Updated on 25-MAR-2003 to								
CC	correct PF field.)								
XX									
SQ	Sequence 5 AA;								
Query Match 91.3%; Score 21; DB 2; Length 5;									
Best Local Similarity 100.0%; Pred. No. 1.8e+06;									
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	3	PHG 5							
Db	2	PHG 4							
RESULT 4									
AAR69933									
ID	AAR69933 standard; peptide; 5 AA.								
XX									
AC	AAR69933;								
XX									
DT	25-MAR-2003 (revised)								
DT	18-OCT-1995 (first entry)								
XX									
DE	Pentameric mimotope 59 used to obtain highly specific antibodies.								
XX									
KW	mimotope; antibody; production; high specificity; detection; immunoassay;								
KW	high performance liquid chromatography.								
XX									
OS	Synthetic.								
XX									
FH	Key Location/Qualifiers								
FT	Modified-site 1								
FT	/note= "Acetyl-His"								

XX (KAUV/) KAUVAR L M.
PA Kauvar LM;
XX WPI; 1996-361955/36.
XX
DR Identifying candidate drugs that bind a specific receptor by competitive
XX reaction with panel of mimetope(s) - useful in rational drug design.
PT
XX Example 3; Fig 3; 27pp; English.
PS
XX The present peptide is a member of a panel of 88 pentapeptide mimotopes
CC designed on the basis of decreasing hydrophobicity, and periodic
CC variation of hydrophobic moment. The panel was synthesised using the
CC method of Geysen, H. M., et al, Proc. Natl. Acad. Sci. USA (1984), which
CC uses lots of 96 pins; the remaining 8 polyethylene pins being controls.
CC The mimotopes were then mixed, 125-I labelled and tested with individual
CC members of a basal antibody (Ab) repertoire. Nearly uniform binding to
CC all members was found. The test was then repeated with the addition of a
CC defined amt. of analyte. A small number showed greatly increased
CC labelling, these Ab representing the successful result of an initial
CC screen for those that pref. bind analyte. The above is an example of a
CC claimed method for identifying members of a panel of candidate drugs,
CC that bind to a receptor having a known ligand. It is useful in rational
CC drug design, e.g. selection of monoclonal Ab for treating individual
CC tumours, and for the prodn. of immunological reagents for any analyte,
CC including those not normally detectable by immunoassay, e.g trace
CC contaminants in soil, air or water. (Updated on 25-MAR-2003 to correct PF
CC field.)
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
2 PHG 4

RESULT 7
AAW78934
ID AAW78934 standard; peptide; 5 AA.
XX
AC AAW78934;
XX
DT 20-NOV-1998 (first entry)
XX
DE Antidepressant MIF based peptide #64.
XX
KW Antidepressant; depression; tripeptide MIF; hormone; bovine; rat;
KW melanocyte stimulating inhibitory factor; psychiatric disorder;
KW neurochemical; central nervous system.
XX
OS Synthetic.
OS Bos taurus.
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "4F-Phe"
FT Modified-site 2 /label= 4Hyp
FT Modified-site 5 /note= "4-hydroxyproline"
FT Modified-site /note= "amidated"
XX
PN US5767083-A.
XX
PD 16-JUN-1998.
XX

PF 02-MAY-1995; 95US-00432651.
XX
PR 04-MAY-1994; 94US-00238089.
XX
PA (INNA-) INNAPHARMA INC.
XX
PI Hlavka JJ, Noble JF, Abajian HB;
XX WPI; 1998-361753/31.
DR
XX New peptides based on structure of MIF - useful for treatment of
PT depression.
PT
PS Claim 14; Col 99; 56pp; English.
XX
CC Peptides given in AAW78922 to AAW79000 are based on the tripeptide MIF
CC (melanocyte stimulating inhibitory factor), which was initially isolated
CC from bovine and rat hypothalamic extracts. The peptides from the present
CC invention are useful for treating depression, optionally together with
CC other antidepressants e.g. fluoxetine, amitriptyline or sertraline. The
CC peptides avoid the serious side effects of the prior art tricyclics, MAO
CC oxidase inhibitors and second generation antidepressants. They are also
CC lipid soluble and easily cross the blood-brain barrier, but are not so
CC easily metabolised as the parent MIF (the tripeptide Pro-Leu-Gly-NH2) and
CC so allows a lower dosage
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
2 PHG 4

RESULT 8
AAW23419
ID AAY23419 standard; peptide; 5 AA.
XX
AC AAY23419;
XX
DT 02-SEP-1999 (first entry)
XX
DE V beta 6 clone found in MS patients after vaccination with TCR.
XX
KW Vaccine; T cell receptor; TCR; T cell; V beta 6.2/3; V beta 6/5;
KW V beta 6.7; V beta 2; V beta 5/1; V beta 7; V beta 13; V beta 6;
KW multiple sclerosis.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9927957-A1.
XX
PD 10-JUN-1999.
XX
PF 03-DEC-1997; 97WO-US023147.
XX
PR 03-DEC-1997; 97WO-US023147.
XX
PA (IMMU-) IMMUNE RESPONSE CORP.
PA (KIMM-) KIMMEL CANCER CENT SIDNEY.
XX
PI Brostoff SW, Wilson DB, Smith LR, Gold DP, Carlo DJ;
XX WPI; 1999-404801/34.
DR
XX T0 cell receptor peptide-derived vaccines.
PS Example 11; Page 85; 104pp; English.
XX

CC The specification describes vaccines which comprise immunologically
CC effective amounts of T cell receptor (TCR) peptides. The TCRs are present
CC on the surface of T cells. The TCRs are chosen from V beta 6.2/3, V beta
CC 6/5, V beta 6.7, V beta 2, V beta 5/1, V beta 7 or V beta 13. The V beta
CC TCR peptide-based vaccines are useful for prevention or treatment of
CC multiple sclerosis (MS). The presence of V beta 6.7 appears to be
CC particularly associated with multiple sclerosis and can be used to
CC determine an individual's susceptibility to multiple sclerosis.
CC Vaccinating, rather than passively administering heterologous antibodies,
CC allows the host's own immune system to mobilize and suppress auto
CC aggressive T cells. Therefore, the suppression is persistent and may
CC involve any and all immunological mechanisms in effecting that
CC suppression. Such a multi-faceted response is more effective than the uni
CC -dimensional suppression achieved by passive administration of monoclonal
CC antibodies or extant-derived regulatory T cell clones. AAY23387-Y23480
CC represent peptides derived from TCR V beta 6 clones found in the
CC cerebrospinal fluid (CSF) of MS patients, after vaccination with V beta 6
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 9
ABB79810
ID ABB79810 standard; peptide; 5 AA.

XX ABB79810;

AC ABB79810;

XX 25-NOV-2002 (first entry)

DT 25-NOV-2002 (first entry)

XX Histidine proline rich glycoprotein pentapeptide.

XX Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW neutropic; neuroprotective; antiparkinsonian.
XX
OS Synthetic.
XX WO200264621-A2.
PN
XX 22-AUG-2002.

XX 14-FEB-2002; 2002WO-US004336.

PF 14-FEB-2002; 2002WO-US004336.

PR 14-FEB-2001; 2001US-0268370P.

XX (ATTE-) ATTENUON LLC.

PA Donate F, Harris S, Plunkett ML, Mazar AP;

XX WPI; 2002-666989/71.

XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 2; Page 67; 82pp; English.

XX The present sequence is a specific example of claimed anti-angiogenic
CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
CC peptides comprise: the histidine-proline-rich (H/P) domain of human
CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic

CC sequence given in ABB79808, such as the present peptide, or its variant
CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
CC N- or C-terminus. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX Sequence 5 AA;

Query Match 91.3%; Score 21; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 3 PHG 5

RESULT 10
ABB79808
ID ABB79808 standard; peptide; 5 AA.

XX ABB79808;

XX 25-NOV-2002 (first entry)

DT Histidine proline rich glycoprotein pentapeptide.

XX Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW neutropic; neuroprotective; antiparkinsonian.
XX
OS Synthetic.

XX Key Location/Qualifiers
FH Misc-difference 1 /label= His, Pro
FT Misc-difference 2 /label= His, Pro
XX WO200264621-A2.

XX PD 22-AUG-2002.

XX PF 14-FEB-2002; 2002WO-US004336.

XX PR 14-FEB-2001; 2001US-0268370P.

XX PA (ATTE-) ATTENUON LLC.

XX PI Donate F, Harris S, Plunkett ML, Mazar AP;

XX XX WPI; 2002-666989/71.

DR P-PSDB; ABN84910.

XX PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for

PT diagnosing or treating diseases associated with undesired cell migration,

PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX XX

PS Claim 1; Page 67; 82pp; English.

XX XX

CC The present sequence is a generic sequence (see also ABB79809-11) of

CC isolated anti-angiogenic pentapeptides of the invention. Claimed anti-

CC angiogenic polypeptides or peptides comprise: the histidine proline-rich

CC (H/P) domain of human histidine proline rich glycoprotein (HPRG, see

CC ABB79806); an H/P domain of rabbit HPRG (see ABB79807); a variant of

CC these that is capable of inhibiting angiogenesis, endothelial cell

CC proliferation or endothelial tube formation in vitro or in vivo; and the

CC present peptide, or its variant having an additional 1 to 4 amino acids

CC comprising His, Pro or Gly at its N- or C-terminus. Also claimed are:

CC chemically synthesised or recombinantly produced peptide multimers; a

CC diagnostically or therapeutically labeled anti-angiogenic polypeptide,

CC peptide or peptide multimer; a diagnostically useful HPRG-related

CC composition, comprising the diagnostically labeled polypeptide, peptide

CC or peptide multimer and a carrier; an antibody specific for an epitope of

CC HPRG that is present in the H/P domain of human or rabbit HPRG, and which

CC binds to HPRG or to any of the domains in a way which inhibits the anti-

CC angiogenic activity of HPRG or the domain, or an antigen-binding fragment

CC of the antibody; a method for inhibiting cell migration, cell invasion,

CC cell proliferation or angiogenesis, or for inducing apoptosis; a method

CC for treating a subject having a disease or condition associated with

CC undesired cell migration, invasion, proliferation, or angiogenesis; a

CC method of stimulating or inhibiting angiogenesis in a subject; a method

CC of detecting the presence of HPRG or its cleavage product or its peptide

CC in a biological sample; isolated nucleic acids encoding the polypeptide,

CC peptide or peptide multimer; an expression vector; transformed or

CC transfected cells; a method of providing to a cell, tissue or organ an

CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its

CC pentapeptide, or the peptide multimer; an affinity ligand useful for

CC binding to, or isolating, an HPRG-binding molecule or cells expressing

CC the binding molecule, comprising the polypeptide, peptide or peptide

CC multimer, immobilised to a solid support or carrier; and a method of

CC isolating HPRG-binding molecule, or isolating or enriching cells

CC expressing HPRG-binding site or receptor, from a complex mixture. The

CC compositions and methods are useful in diagnosing or treating a disease

CC or condition associated with undesired cell migration, invasion,

CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic

CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or

CC fractures. HPRG is especially useful in inhibiting the growth of primary

CC tumours or metastases, and may also be used in treating neurodegenerative

CC diseases like Alzheimer's or Parkinson's disease. The antibodies are

CC stimulators of angiogenesis and are useful for promoting

CC neovascularization in pertinent disease states, and in various

CC immunoassays

XX SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 5; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5

Db 3 PHG 5

RESULT 11

ABB79809

ID ABB79809 standard; peptide; 5 AA.

XX AC ABB79809;

XX DT 25-NOV-2002 (first entry)

XX XX Histidine proline rich glycoprotein pentapeptide.

DE Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;

XX KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;

KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;

KW nootropic; neuroprotective; antiparkinsonian.

XX OS Synthetic.

XX XX WO200264621-A2.

PN 22-AUG-2002.

PD 14-FEB-2002; 2002WO-US004336.

XX PF 14-FEB-2001; 2001US-0268370P.

PR (ATTE-) ATTENUON LLC.

PA Donate F, Harris S, Plunkett ML, Mazar AP;

XX PI WPI; 2002-666989/71.

XX DR P-PSDB; ABN84910.

XX XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for

PT diagnosing or treating diseases associated with undesired cell migration,

PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX XX

PS Claim 2; Page 67; 82pp; English.

XX XX

CC The present sequence is a specific example of claimed anti-angiogenic

CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or

CC peptides comprise: the histidine-proline-rich (H/P) domain of human

CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain

CC of rabbit HPRG (see ABB79807); a variant of these that is capable of

CC inhibiting angiogenesis, endothelial cell proliferation or endothelial

CC tube formation in vitro or in vivo; or a pentapeptide having the generic

CC sequence given in ABB79808, such as the present peptide, or its variant

CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its

CC N- or C-terminus. Also claimed are: chemically synthesised or

CC recombinantly produced peptide multimers; a diagnostically or

CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide

CC multimer; a diagnostically useful HPRG-related composition, comprising

CC the diagnostically labeled polypeptide, peptide or peptide multimer and a

CC carrier; an antibody specific for an epitope of HPRG that is present in

CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any

CC of the domains in a way which inhibits the anti-angiogenic activity of

CC HPRG or the domain, or an antigen-binding fragment of the antibody; a

CC method for inhibiting cell migration, cell invasion, cell proliferation

CC or angiogenesis, or for inducing apoptosis; a method for treating a

CC subject having a disease or condition associated with undesired cell

CC migration, invasion, proliferation, or angiogenesis; a method of

CC stimulating or inhibiting angiogenesis in a subject; a method of

CC detecting the presence of HPRG or its cleavage product or its peptide in

CC a biological sample; isolated nucleic acids encoding the polypeptide,

CC peptide or peptide multimer; an expression vector; transformed or

CC transfected cells; a method of providing to a cell, tissue or organ an

CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its

CC pentapeptide, or the peptide multimer; an affinity ligand useful for

CC binding to, or isolating, an HPRG-binding molecule or cells expressing

CC the binding molecule, comprising the polypeptide, peptide or peptide

CC multimer, immobilised to a solid support or carrier; and a method of

CC isolating HPRG-binding molecule, or isolating or enriching cells

CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db |||
3 PHG 5

RESULT 12
ABB79811
ID ABB79811 standard; peptide; 5 AA.
XX
AC ABB79811;
XX
DT 25-NOV-2002 (first entry)
XX
DE Histidine proline rich glycoprotein pentapeptide.
XX
KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW neurotropic; neuroprotective; antiparkinsonian.
XX
OS Synthetic.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 2; Page 67; 82pp; English.
XX
CC The present sequence is a specific example of claimed anti-angiogenic
CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
CC peptides comprise: the histidine-proline-rich (H/P) domain of human
CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808, such as the present peptide, or its variant
CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
CC N- or C-terminus. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising

CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db |||
3 PHG 5

RESULT 13
AAG80338
ID AAG80338 standard; peptide; 5 AA.
XX
AC AAG80338;
XX
DT 20-FEB-2002 (first entry)
XX
DE Enzyme cleavable peptide SEQ ID 18.
XX
KW Antineoplastic; drug; cytostatic; chemotherapeutic agent; matrixin;
KW cancer; leukaemia; lymphoma; carcinoma; sarcoma; melanoma.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Acetylated residue"
FT Modified-site 4 /label= OTHER
FT Modified-site 5 /note= "Homophenylalanine, Hof"
FT /label= OTHER
FT /note= "Residue is conjugated to Dox, where Dox is the
FT anthracycline doxorubicin"
XX
PN WO200168145-A2.
XX
PD 20-SEP-2001.
XX

PF 15-MAR-2001; 2001WO-US008589.
XX
PR 15-MAR-2000; 2000US-0189387P.
XX
PA (DUPO) DUPONT PHARM CO.
XX
PI Copeland RA, Albright CF, Combs AP, Dowling RL, Graciani NR;
PI Han W, Higley CA, Huang PS, Yue EW, Dimeo SV;
XX
DR WPI; 2002-010546/01.
XX
XX New compounds comprising enzyme cleavable peptide conjugated to
PT antineoplastic agent used as chemotherapeutic agents in targeted
PT treatment of cancers.
XX
XX Claim 33; Page 112; 203pp; English.
PS
XX This invention relates to novel compounds (I) comprising antineoplastic
CC agents conjugated to enzyme cleavable peptides comprising the amino acid
CC recognition sequence of a membrane bound and/or cell-secreted peptidase.
CC The compounds of the invention have cytostatic activity. The products of
CC the invention can be used as chemotherapeutic agents, preferably in the
CC presence of a matrixin, used for targeted treatment of cancers,
CC particularly leukaemia, lymphoma, carcinoma, sarcoma and melanoma. (I)
CC are stable to non-specific enzymatic degradation in the blood stream and
CC reduce damage to healthy, non-target tissue, so that the toxicity of (I)
CC is reduced, compared to its non-conjugated form. This sequence represents
CC an enzyme cleavable peptide described in the method of the invention
XX
XX Sequence 5 AA;
SQ

Query Match 91.3%; Score 21; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 1 PHG 3

RESULT 14
AAO15016
ID AAO15016 standard; peptide; 5 AA.
XX
AC AAO15016;
XX
DT 16-AUG-2002 (first entry)
XX
DE Mutant VL CDR3 peptide (C10).
XX
KW Ribosome display system; specific binding pair; sbp; VL CDR3; C10;
KW scFv antibody generation; GPI-linked cell surface receptor; mutein.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO200175097-A2.
XX
PD 11-OCT-2001.
XX
PF 26-MAR-2001; 2001WO-GB001319.
XX
PR 31-MAR-2000; 2000US-0193802P.
XX
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Osbourn J, Holet T;
XX
DR WPI; 2002-010788/01.
XX
XX Obtaining specific binding pair member (I) that binds a complementary
PT specific binding pair member of encapsitating specific binding pair
PT member/ribosome complexes in a viral coat, comprises the use of a

PT ribosome display system.
XX
XX Example 9; Page 49; 6lpp; English.
XX
CC The invention comprises a method (ribosome display system) for obtaining
CC a specific binding pair (sbp) member that binds a complementary sbp
CC member of interest. The method involves incubating mRNA molecules
CC encoding an sbp and lacking an in-frame stop codon, allowing ribosome
CC translation of the mRNA to produce the encoded sbp member, forming
CC complexes comprising ribosome, mRNA, and encoded sbp member displayed on
CC the ribosome. The ribosome display system is useful for the selection of
CC an sbp member able to bind a complementary sbp member. The present amino
CC acid sequence represents a mutant VL CDR3 peptide
XX
XX Sequence 5 AA;
SQ

Query Match 91.3%; Score 21; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 2 PHG 4

RESULT 15
ADH10419
ID ADH10419 standard; peptide; 5 AA.
XX
AC ADH10419;
XX
DT 11-MAR-2004 (first entry)
XX
DE HPRG protein H/P rich domain mutated repeat fragment.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Synthetic.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI McCrae K, Donate F, Juarez J, Mazar AP;
XX
XX WPI; 2004-090604/09.
DR
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Example 5; SEQ ID NO 33; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the

CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a mutated HPRG
CC protein His-Pro (H/P) rich domain repeat fragment.
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 3 PHG 5

RESULT 16
ADH10422

ID ADH10422 standard; peptide; 5 AA.
XX
AC ADH10422;
XX
DT 11-MAR-2004 (first entry)
XX
DE HPRG protein H/P rich domain mutated repeat fragment.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Synthetic.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX
DR WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Example 5; SEQ ID NO 36; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,

CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a mutated HPRG
CC protein His-Pro (H/P) rich domain repeat fragment.
XX
SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 3 PHG 5

RESULT 17
ADH10417

ID ADH10417 standard; peptide; 5 AA.
XX
AC ADH10417;
XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein H/P rich domain repeat fragment.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Oryctolagus cuniculus.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX
DR WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Example 5; SEQ ID NO 31; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related

CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 3 PHG 5

RESULT 18

ADH10415
ID ADH10415 standard; peptide; 5 AA.

XX ADH10415;

DT 11-MAR-2004 (first entry)

DE Rabbit HPRG protein H/P rich domain repeat fragment.

XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.

XX Oryctolagus cuniculus.

XX WO2003077872-A2.

PD 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

PF 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

XX Mccrae K, Donate F, Juarez J, Mazar AP;

DR WPI; 2004-090604/09.

XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.

XX Example 5; SEQ ID NO 29; 117pp; English.
PS The invention relates to an isolated tropomyosin (Tpm)-related
XX antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 3 PHG 5

RESULT 19

ADH10416
ID ADH10416 standard; peptide; 5 AA.

XX ADH10416;

DT 11-MAR-2004 (first entry)

DE Rabbit HPRG protein H/P rich domain repeat fragment.

XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.

XX Oryctolagus cuniculus.

XX WO2003077872-A2.

PD 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

PR 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

XX Mccrae K, Donate F, Juarez J, Mazar AP;

DR WPI; 2004-090604/09.

PT	New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT	inhibiting endothelial cell migration, invasion, proliferation or
PT	angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT	cancer.
XX	
PS	Example 5; SEQ ID NO 30; 117pp; English.
XX	
CC	The invention relates to an isolated tropomyosin (Tpm)-related
CC	antiangiogenic receptor polypeptide or peptide, which is a fragment
CC	of a full-length native Tpm protein expressed on the surface of
CC	endothelial cells, or a variant of the fragment. It has a molecular mass
CC	of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC	an internal fragment of a native Tpm isoform which is a binding site for
CC	antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC	polypeptide, peptide or variant has substantially the same biochemical
CC	activity of binding to the antiangiogenic polypeptide agents, as does the
CC	native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC	binds to the isolated polypeptide or peptide is human histidine-proline
CC	rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC	rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC	homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC	human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC	antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC	domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC	antibodies and compositions are useful for inhibiting endothelial cell
CC	migration, invasion, proliferation or angiogenesis, for inducing
CC	endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC	retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC	psoriasis, or scleroderma. The antibody may be also used for detecting
CC	the presence of a Tpm polypeptide or peptide in a biological sample, for
CC	promoting wound healing, or for treating diseases or conditions in which
CC	increased angiogenesis is desired, e.g. coronary artery disease or
CC	peripheral artery disease. The present sequence represents a repeat
CC	fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.
XX	
SQ	Sequence 5 AA;
	Query Match 91.3%; Score 21; DB 8; Length 5;
	Best Local Similarity 100.0%; Pred. No. 1.8e+06;
	Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	3 PHG 5
Db	3 PHG 5
RESULT 20	
ADH10418	
ID	ADH10418 standard; peptide; 5 AA.
XX	
AC	ADH10418;
XX	
DT	11-MAR-2004 (first entry)
XX	
DE	HPRG protein H/P rich domain consensus repeat fragment.
XX	
KW	Tpm; tropomyosin; antiangiogenic receptor;
KW	histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW	ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW	antipsoriatic; dermatological; cardiac; vasotropic; vulnery;
KW	angiogenesis; gene therapy; rabbit.
XX	
OS	Synthetic.
OS	Oryctolagus cuniculus.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 1 /label= His or Pro
FT	Misc-difference 2 /label= His or Pro
FT	
FT	
XX	
PN	WO2003077872-A2.
XX	

XX OS Ovis aries.
XX PN DE10236711-A1.
XX PD 26-FEB-2004.
XX PF 09-AUG-2002; 2002DE-01036711.
XX PR 09-AUG-2002; 2002DE-01036711.
XX PA (UYHO-) UNIV HOHENHEIM.
XX PI Geldermann H, Preuss S, Han Y;
XX DR WPI; 2004-215730/21.
XX PT Typing genes that contain polymorphic microsatellite loci, useful for
PT identifying predisposition to disease, by amplification and determining
PT length of amplicons.
XX PS Disclosure; Page 9; 64pp; German.
XX CC The invention describes a method of typing (M1) a gene (I) that has one
CC or more polymorphic microsatellite loci (PML). The method comprises: PCR
CC amplification of at least one DNA region of (I) that includes PML, using
CC as template a DNA sample containing at least one segment of (I); and
CC determining the length of the resulting amplicon(s). Also described are:
CC a method of determining (M2) microsatellite markers (NM) for
CC predisposition to a disease, associated with a gene that includes one or
CC more PML; and prediagnosis (M3) of diseases associated with gene that
CC include PML. The method is used to identify microsatellite markers, in a
CC disease-related gene, that are associated with a predisposition to
CC diseases and for prediagnosis of such diseases, especially prion diseases
CC but also cystic fibrosis, malignant hyperthermia syndrome in pigs and
CC metabolic diseases; also to type genes that encode milk proteins,
CC hormones or transcription factors. The method is simpler, quicker and
CC particularly less expensive than known methods based on sequencing. This
CC is the amino acid sequence of a sheep prion protein motif.
XX Sequence 5 AA;
SQ
Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db |||
1 PHG 3
RESULT 22
ADR03627
ID ADR03627 standard; peptide; 5 AA.
XX
AC ADR03627;
XX 21-OCT-2004 (first entry)
XX
DE E. coli membrane protein ligand #25, SEQ ID 64.
XX
KW Ligand; identification; drug target; cancer; cardiovascular disease;
KW autoimmune disease; infection; inflammatory disease;
KW central nervous system disorder; metabolic disease; endocrine disease;
KW membrane protein.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Xaa = T(Sa), Sialic acid threonine lactam"
XX
PN WO2004062553-A2.

XX 29-JUL-2004.
XX 16-JAN-2004; 2004WO-DK000023.
XX 16-JAN-2003; 2003US-00346737.
XX PR 19-MAY-2003; 2003DK-00000749.
XX (CARL-) CARLSBERG AS.
XX Hilaire SPM, Yin H, Surve S, Wenckens M;
XX WPI; 2004-593147/57.
XX Identifying members of protein-ligand binding pair useful as drug target,
PT by incubating immobilized ligand library with differentially labeled
PT protein, detecting ligand-protein pair, identifying protein and ligand of
PT pair.
XX Example 33; SEQ ID NO 64; 172pp; English.
XX The present invention relates to a method (M1) for identifying specific
CC members of an unknown protein-ligand binding pair. The method involves
CC synthesizing an immobilized ligand library, incubating the library with
CC differentially labelled protein mixtures, detecting the immobilized
CC ligand-protein pair and identifying the protein and ligand of the pair.
CC The identified ligand and protein are specific members of an unknown
CC differential ligand-protein binding pair. (M1) is rapid and efficient,
CC highly accurate and is able to take a large number of unknown ligands
CC and/or unknown proteins. (M1) is used for identifying proteins suitable
CC as drug targets, to identify one or more drugs for treating clinical
CC conditions such as cancer, cardiovascular diseases, autoimmune diseases,
CC infections, inflammatory diseases, central nervous system disorders,
CC metabolic disease and endocrine diseases. The present sequence is a
CC peptide ligand used to illustrate the invention.
XX Sequence 5 AA;
SQ
Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db |||
3 PHG 5
RESULT 23
AAW75370
ID AAW75370 standard; peptide; 6 AA.
XX
AC AAW75370;
XX 02-FEB-1999 (first entry)
XX
DE Hexapeptide #10 binds immobilised anti-Ad5 fibre head MAb 7A2.7.
XX
KW Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;
KW viral surface protein; hexapeptide expression library; adenovirus;
KW major histocompatibility complex; MHC; fibronectin; gene therapy;
KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.
XX Synthetic.
XX
PN FR2758821-A1.
XX 31-JUL-1998.
XX
PF 30-JAN-1997; 97FR-00001005.
XX
PR 30-JAN-1997; 97FR-00001005.
XX (CNRS) CNRS CENT NAT RECH SCI.
PA

CC The present invention describes a purified elastase variant (I) with an
CC amino acid sequence different from that of a precursor elastase, the
CC difference comprising a substitution of an active site histidine residue
CC corresponding to residue 43 in human neutrophil elastase with a different
CC amino acid residue so that (I) has substrate specificity substantially
CC different from the precursor elastase. (I) has cytotstatic activity, and
CC can be used in antibody-directed enzyme activated prodrug therapy. The
CC elastase variant can be used to cleave a particular substrate, especially
CC those containing histidine residues at the substrate site. Site-specific
CC proteolysis is useful in therapeutic applications, e.g. for antibody-
CC directed enzyme activated prodrug therapy (ADEPT). AAC88022, AAC88023 and
CC AAB55432 to AAB55526 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 2 PHG 4

RESULT 26
ABG31535
ID ABG31535 standard; peptide; 6 AA.

XX
AC ABG31535;

XX
DT 05-NOV-2002 (first entry)

XX
DE Camptothecin peptide conjugate #24.

XX
KW Camptothecin; integrin receptor antagonist; cytotstatic agent;
KW tumour cell; metallo matrix protease; MMP; carcinomatous disorder;
KW integrin alpha_vbeta_3 receptor antagonist; tumour growth inhibitor;
KW tumour; integrin; camptothecin-bis-trifluoroacetate.

OS Synthetic.

XX
FH Key Location/Qualifiers

FT Modified-site 6

FT /label= OTHER
FT /note= "Bound to Camptothecin bis-trifluoroacetate at
FT position 20-O"

XX
PN EP1219305-A1.

XX
PD 03-JUL-2002.

XX
PF 27-DEC-2000; 2000EP-00128401.

XX
PR 27-DEC-2000; 2000EP-00128401.

XX
PA (FARB) BAYER AG.

XX
PI Lerchen H, Baumgarten J, Lockhoff O, Albers M, Schoop A;

XX
DR WPI; 2002-576993/62.

XX
PT New conjugates of integrin receptor antagonist and a cytotstatic agent
PT with specific cleavable linking unit useful in the treatment of cancer.

XX
PS Example; Page 61; 127pp; English.

XX
CC This invention relates to conjugates of integrin receptor antagonist and
CC a cytotstatic agent with cleavable linking unit that are selectively
CC cleaved by tumour cell metallo matrix proteases (MMPs). The conjugates of
CC the invention may have cytotstatic activity and may be used as an integrin
CC alpha_vbeta_3 receptor antagonist or a tumour growth inhibitor. The
CC conjugates of the invention may be used in the production of a medicament

CC for the treatment of carcinomatous disorders. Compounds of the invention
CC containing the conjugates of the invention exhibit a selective and tumour
CC -specific action as a result of linkage to alpha_vbeta_3 integrin
CC antagonists via linking units which can be selectively cleaved by enzymes
CC such as metallo matrix proteases, i.e. enzymes found in tumour tissue.
CC The linking units can also maintain the serum stability of the conjugate
CC of cytotstatic and alpha_vbeta_3 integrin antagonist, and at the same
CC time, show the desired intracellular action within tumour cells as a
CC result of its specific enzymatic and hydrolytic cleavability with release
CC of the cytotstatic. The present sequence represents a camptothecin-bis-
CC trifluoroacetate conjugated peptide used in an example of the method of
CC the invention
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 1 PHG 3

RESULT 27
AAE26391
ID AAE26391 standard; peptide; 6 AA.

XX
AC AAE26391;

XX
DT 13-DEC-2002 (first entry)

XX
DE Human GPR10 mutant C-terminal peptide, dell6.

XX
KW Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;
KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
KW anticonvulsant; mutant; mutein.

OS Homo sapiens.

OS Synthetic.

XX
PN US2002037533-A1.

XX
PD 28-MAR-2002.

XX
PF 17-AUG-2001; 2001US-00932161.

XX
PR 28-APR-2000; 2000US-00560915.

XX
PA (CIVE/) CIVELLI O.
PA (LINS/) LIN S.

XX
PI Civelli O, Lin S;

XX
DR WPI; 2002-403931/43.

XX
PT Screening for compounds useful for promoting wakefulness or sleep, and
PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
PT apnea, comprises administering a prolactin releasing peptide agonist or
PT antagonist.

XX
PS Example 2; Page 22; 35pp; English.

XX
CC The present invention relates to a method of screening for compounds for
CC promoting wakefulness or sleep in a mammal. The method involves
CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
CC agonist or antagonist respectively and determining the ability of the
CC compound to promote wakefulness or sleep. The compounds identified from
CC the method are used in the therapy of epilepsy and other diseases
CC associated with absence seizures and in promoting wakefulness and sleep
CC in individuals having sleep disorders such as insomnia and narcolepsy.
CC PrRP receptor agonists may be used to treat common disorders which lead

CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
CC promoting sleep and for treating insomnia such as adjustment sleep
CC disorder and psychophysiologic insomnia. The present sequence is human
CC GPR10 mutant C-terminal peptide
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 1 PHG 3

RESULT 28
ABP53526
ID ABP53526 standard; peptide; 6 AA.
XX
AC ABP53526;
XX
DT 12-DEC-2002 (first entry)
XX
DE Camptothecin peptide conjugate II.22.
XX
KW Camptothecin peptide conjugate; cytostatic; glycoconjugate; tumour;
KW specifically cleavable peptidic linking unit; cancer.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminally modified with 20-O-"
FT Modified-site 6
FT /note= "C-terminally modified with -camptothecin TFA"
XX
XX EP1219634-A1.
XX
PD 03-JUL-2002.
XX
PF 27-DEC-2000; 2000EP-00128402.
XX
PR 27-DEC-2000; 2000EP-00128402.
XX
PA (FARB) BAYER AG.
XX
PI Lerchen H, Baumgarten J, Lockhoff O;
XX
DR WPI; 2002-629644/68.
XX
PT Cytostatic-glycoconjugates useful for treating cancer have specifically
PT cleavable peptide linking units so as to deliver drug to tumor only.
XX
PS Example; Page 17; 46pp; English.
XX
CC The present invention describes a conjugate (I) and its salts. (I) has
CC the formula CT-LI-Sp1-Sp2-K, where: CT = cytotoxic radical or radical of
CC a cytostatic derivative optionally also carrying a OH, COOH or NH2 group;
CC LI = linker comprising 5 to 8 amino acids each optionally carrying
CC protecting groups; Sp1 = absent, CO or CS; Sp2 = optionally substituted
CC arylene or alkylene; and K = unsubstituted or regioselectively modified
CC carbohydrate radical. Also described: (1) a process for preparation of
CC the conjugates; (2) a medicament comprising the conjugate. (I) has
CC cytostatic activity. The conjugates are useful for treating cancer. The
CC conjugate is cleaved by enzymes found especially in tumour tissue to
CC release the cytotoxic drug. As the drug is not active until after
CC cleavage it will not cause harm to non-tumour proliferating cells,
CC therefore reducing the side-effects associated with the use of such
CC drugs. The present sequence represents a camptothecin peptide conjugate,
CC which is used in an example from the present invention
XX

SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 1 PHG 3

RESULT 29
ADN12399
ID ADN12399 standard; peptide; 6 AA.
XX
AC ADN12399;
XX
DT 17-JUN-2004 (first entry)
XX
DE Peptide #1 of the invention.
XX
KW ubiquitin-specific protease; neuroplasticity; neurodegeneration.
XX
OS Rattus norvegicus.
XX
PN WO2003038097-A1.
XX
PD 08-MAY-2003.
XX
PF 29-OCT-2002; 2002WO-JP011231.
XX
PR 29-OCT-2001; 2001JP-00330339.
XX
PA (EISA) EISAI CO LTD.
XX
PI Suzuki T, Tian QB;
XX
DR WPI; 2003-430529/40.
XX
PT Brain-occurring ubiquitin-specific protease and encoded DNA, applicable
PT in studying molecular mechanism of expression neuroplasticity and
PT pathosis of neurodegeneration.
XX
PS Disclosure; SEQ ID NO 3; 33pp; Japanese.
XX
CC The present invention relates to a protein comprising a 1036 residue
CC amino acid sequence having ubiquitin-specific protease activity. The
CC protein and its encoded DNA are applicable in studying the molecular
CC mechanism of expression neuroplasticity and pathosis of
CC neurodegeneration. The present sequence represents the a peptide of the
CC invention.
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 4 PHG 6

RESULT 30
ADO37172
ID ADO37172 standard; peptide; 6 AA.
XX
AC ADO37172;
XX
DT 29-JUL-2004 (first entry)
XX
DE Binding partner polypeptide of the invention SEQ ID NO:232.
XX

XX WPI; 2004-376185/35.

DR

XX

PT Evenly distributing tags among members of a starting library, useful in

PT developing pharmaceuticals and diagnostics, comprises dividing the

PT starting library into sub libraries and attaching a tag to members of

PT each sub library.

XX

XX Claim 141; SEQ ID NO 957; 510pp; English.

PS

XX

XX The invention relates to novel methods for producing polypeptide-tagged

CC collections and capture systems containing the tagged polypeptides. The

CC method is useful for evenly distributing tags among members of a starting

CC library. The system, collection, kits and methods are useful in

CC developing pharmaceuticals and diagnostics. The present sequence is used

CC in the exemplification of the invention.

XX

XX

SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5

Db 2 PHG 4

RESULT 36

ADO28227

ID ADO28227 standard; peptide; 6 AA.

XX

AC ADO28227;

XX

DT 12-AUG-2004 (first entry)

XX

DE Capture system related peptide, SEQ ID 765.

XX

KW Capture system.

XX

XX Synthetic.

OS

XX WO2004042019-A2.

PN

PD 21-MAY-2004.

XX

XX 30-OCT-2003; 2003WO-US034693.

PF

XX 30-OCT-2002; 2002US-0422923P.

PR

PR 30-OCT-2002; 2002US-0423018P.

XX

PA (POIN-) POINTILLISTE INC.

XX

XX Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;

PI

XX WPI; 2004-431543/40.

DR

XX

PT Capturing biological particles, by contacting biological particles with

PT capture system comprising addressed loci, addressed collection of

PT polypeptide tagged molecules, capture agents, and polypeptide tag to

PT which capture agent binds.

XX

PS Disclosure; SEQ ID NO 765; 505pp; English.

XX

CC The present invention relates to a method for the capture and analysis of

CC biological particle using a capture system. The method is useful for

CC capturing biological particles such as cells, portions of cells, cell

CC membranes, viruses, viral capsids, viral particles, bacterial cells,

CC subcellular compartments, organelles and micelles, prokaryotic cells,

CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell

CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral

CC vectors or viral capsids with or without packaged nucleic acid, phage,

CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are

CC cells chosen from immune cells, neurons, cancer cells, bacterial cells

CC and infected cells, subcellular compartment, organelles, viral particles

CC or pathogens. The cells are dendritic cells, T cells, or B cells. The

CC method is also useful for identifying molecules that interact with

CC infectious agents, for profiling the surface of a biological particles,

CC for identifying a modulator of an interaction among proteins in the

CC biological particle, for identifying molecules that modulates the

CC trafficking, activity or functional or structural property in the

CC biological particle, and for mapping epitopes of molecules displayed on

CC the surface of a biological particles. The method is also useful for

CC sorting biological particles, for identifying a receptor on the surface

CC of biological particle that transduces a signal from a polypeptide, and

CC for identifying the molecule that interacts with an apically-localized

CC molecule on a biological particle. The present sequence was used to

CC illustrate the invention.

XX

XX

SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5

Db 3 PHG 5

RESULT 37

ADO28419

ID ADO28419 standard; peptide; 6 AA.

XX

AC ADO28419;

XX

DT 12-AUG-2004 (first entry)

XX

DE Capture system related peptide, SEQ ID 957.

XX

KW Capture system.

XX

XX Synthetic.

OS

XX WO2004042019-A2.

PN

PD 21-MAY-2004.

XX

XX 30-OCT-2003; 2003WO-US034693.

PF

XX 30-OCT-2002; 2002US-0422923P.

PR

PR 30-OCT-2002; 2002US-0423018P.

XX

PA (POIN-) POINTILLISTE INC.

XX

XX Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;

PI

XX WPI; 2004-431543/40.

DR

XX

PT Capturing biological particles, by contacting biological particles with

PT capture system comprising addressed loci, addressed collection of

PT polypeptide tagged molecules, capture agents, and polypeptide tag to

PT which capture agent binds.

XX

PS Disclosure; SEQ ID NO 957; 505pp; English.

XX

CC The present invention relates to a method for the capture and analysis of

CC biological particle using a capture system. The method is useful for

CC capturing biological particles such as cells, portions of cells, cell

CC membranes, viruses, viral capsids, viral particles, bacterial cells,

CC subcellular compartments, organelles and micelles, prokaryotic cells,

CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell

CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral

CC vectors or viral capsids with or without packaged nucleic acid, phage,

CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
CC and infected cells, subcellular compartment, organelles, viral particles
CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
CC method is also useful for identifying molecules that interact with
CC infectious agents, for profiling the surface of a biological particles,
CC for identifying a modulator of an interaction among proteins in the
CC biological particle, for identifying molecules that modulates the
CC trafficking, activity or functional or structural property in the
CC biological particle, and for mapping epitopes of molecules displayed on
CC the surface of a biological particles. The method is also useful for
CC sorting biological particles, for identifying a receptor on the surface
CC of biological particle that transduces a signal from a polypeptide, and
CC for identifying the molecule that interacts with an apically-localized
CC molecule on a biological particle. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 2 PHG 4

RESULT 38
ADO28426
ID ADO28426 standard; peptide; 6 AA.
XX
AC ADO28426;
XX
DT 12-AUG-2004 (first entry)
XX
DE Capture system related peptide, SEQ ID 964.
XX
KW Capture system.
XX
OS Synthetic.
XX
PN WO2004042019-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034693.
XX
PR 30-OCT-2002; 2002US-0422923P.
PR 30-OCT-2002; 2002US-0423018P.
XX
PA (POIN-) POINTILLISTE INC.
XX
PI Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;
XX
DR WPI; 2004-431543/40.
XX
PT Capturing biological particles, by contacting biological particles with
PT capture system comprising addressed loci, addressed collection of
PT polypeptide tagged molecules, capture agents, and polypeptide tag to
PT which capture agent binds.
XX
PS Disclosure; SEQ ID NO 964; 505pp; English.
XX

CC The present invention relates to a method for the capture and analysis of
CC biological particle using a capture system. The method is useful for
CC capturing biological particles such as cells, portions of cells, cell
CC membranes, viruses, viral capsids, viral particles, bacterial cells,
CC subcellular compartments, organelles and micelles, prokaryotic cells,
CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
CC vectors or viral capsids with or without packaged nucleic acid, phage,
CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
CC and infected cells, subcellular compartment, organelles, viral particles
CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
CC method is also useful for identifying molecules that interact with
CC infectious agents, for profiling the surface of a biological particles,
CC for identifying a modulator of an interaction among proteins in the
CC biological particle, for identifying molecules that modulates the
CC trafficking, activity or functional or structural property in the
CC biological particle, and for mapping epitopes of molecules displayed on
CC the surface of a biological particles. The method is also useful for
CC sorting biological particles, for identifying a receptor on the surface
CC of biological particle that transduces a signal from a polypeptide, and
CC for identifying the molecule that interacts with an apically-localized
CC molecule on a biological particle. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 3 PHG 5

RESULT 39
ADO27694
ID ADO27694 standard; peptide; 6 AA.
XX
AC ADO27694;
XX
DT 12-AUG-2004 (first entry)
XX
DE Capture system related peptide, SEQ ID 232.
XX
KW Capture system.
XX
OS Synthetic.
XX
PN WO2004042019-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034693.
XX
PR 30-OCT-2002; 2002US-0422923P.
PR 30-OCT-2002; 2002US-0423018P.
XX
PA (POIN-) POINTILLISTE INC.
XX
PI Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;
XX
DR WPI; 2004-431543/40.
XX
PT Capturing biological particles, by contacting biological particles with
PT capture system comprising addressed loci, addressed collection of
PT polypeptide tagged molecules, capture agents, and polypeptide tag to
PT which capture agent binds.
XX
PS Disclosure; SEQ ID NO 232; 505pp; English.
XX

CC The present invention relates to a method for the capture and analysis of
CC biological particle using a capture system. The method is useful for
CC capturing biological particles such as cells, portions of cells, cell
CC membranes, viruses, viral capsids, viral particles, bacterial cells,
CC subcellular compartments, organelles and micelles, prokaryotic cells,
CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
CC vectors or viral capsids with or without packaged nucleic acid, phage,
CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
CC and infected cells, subcellular compartment, organelles, viral particles
CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
CC method is also useful for identifying molecules that interact with
CC infectious agents, for profiling the surface of a biological particles,
CC for identifying a modulator of an interaction among proteins in the
CC biological particle, for identifying molecules that modulates the
CC trafficking, activity or functional or structural property in the
CC biological particle, and for mapping epitopes of molecules displayed on
CC the surface of a biological particles. The method is also useful for
CC sorting biological particles, for identifying a receptor on the surface
CC of biological particle that transduces a signal from a polypeptide, and
CC for identifying the molecule that interacts with an apically-localized
CC molecule on a biological particle. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 40
ADO27702
ID ADO27702 standard; peptide; 6 AA.
XX
AC ADO27702;
XX
DT 12-AUG-2004 (first entry)
XX
DE Capture system related peptide, SEQ ID 240.
XX
KW Capture system.
XX
OS Synthetic.
XX
FN WO2004042019-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034693.
XX
PR 30-OCT-2002; 2002US-0422923P.
PR 30-OCT-2002; 2002US-0423018P.
XX
PA (POIN-) POINTILLISTE INC.
XX
PI Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;
XX
DR WPI; 2004-431543/40.
XX
PT Capturing biological particles, by contacting biological particles with
PT capture system comprising addressed loci, addressed collection of
PT polypeptide tagged molecules, capture agents, and polypeptide tag to
PT which capture agent binds.
XX
PS Disclosure; SEQ ID NO 240; 505pp; English.
XX

CC The present invention relates to a method for the capture and analysis of
CC biological particle using a capture system. The method is useful for
CC capturing biological particles such as cells, portions of cells, cell
CC membranes, viruses, viral capsids, viral particles, bacterial cells,
CC subcellular compartments, organelles and micelles, prokaryotic cells,
CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
CC membrane fragments, nuclear membranes, nuclear membrane fragments, viral
CC vectors or viral capsids with or without packaged nucleic acid, phage,
CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
CC and infected cells, subcellular compartment, organelles, viral particles
CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
CC method is also useful for identifying molecules that interact with
CC infectious agents, for profiling the surface of a biological particles,
CC for identifying a modulator of an interaction among proteins in the
CC biological particle, for identifying molecules that modulates the
CC trafficking, activity or functional or structural property in the
CC biological particle, and for mapping epitopes of molecules displayed on
CC the surface of a biological particles. The method is also useful for
CC sorting biological particles, for identifying a receptor on the surface
CC of biological particle that transduces a signal from a polypeptide, and
CC for identifying the molecule that interacts with an apically-localized
CC molecule on a biological particle. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

Search completed: June 15, 2005, 14:15:45
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-7

Perfect score: 23

Sequence: 1 XXPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	91.3	5	1	US-07-729-099-12
2	21	91.3	5	1	US-08-257-392-12
3	21	91.3	5	3	US-08-770-035-12
4	21	91.3	5	4	US-10-122-246A-64
5	21	91.3	5	6	5217869-59
6	21	91.3	5	6	5217869-59
7	21	91.3	6	3	US-09-095-407-3
8	21	91.3	6	3	US-09-560-915-5
9	21	91.3	6	4	US-09-155-613A-71
10	21	91.3	6	5	PCT-US95-04567-11
11	21	91.3	7	3	US-08-776-265-12
12	21	91.3	7	4	US-09-398-184-12
13	21	91.3	8	3	US-08-405-647B-13
14	21	91.3	8	3	US-08-985-499-13
15	21	91.3	8	3	US-09-125-099-2
16	21	91.3	8	4	US-09-239-043D-1149
17	21	91.3	8	4	US-09-239-043D-1209
18	21	91.3	8	4	US-09-112-956-1
19	21	91.3	8	5	PCT-US96-03180-13
20	21	91.3	9	3	US-09-510-738A-102
21	21	91.3	9	3	US-09-510-738A-124
22	21	91.3	9	3	US-09-510-738A-125
23	21	91.3	9	3	US-09-510-738A-140
24	21	91.3	9	3	US-09-510-738A-144
25	21	91.3	9	4	US-09-861-966-102
26	21	91.3	9	4	US-09-861-966-124
27	21	91.3	9	4	US-09-861-966-125

28	21	91.3	9	4	US-09-861-966-140	Sequence 140, App
29	21	91.3	9	4	US-09-861-966-144	Sequence 144, App
30	21	91.3	9	4	US-09-239-043D-1128	Sequence 1128, Ap
31	21	91.3	9	4	US-09-239-043D-1667	Sequence 1667, Ap
32	21	91.3	9	4	US-09-870-089B-13	Sequence 13, Appl
33	21	91.3	9	4	US-09-919-048-102	Sequence 102, App
34	21	91.3	9	4	US-09-919-048-124	Sequence 124, App
35	21	91.3	9	4	US-09-919-048-125	Sequence 125, App
36	21	91.3	9	4	US-09-919-048-140	Sequence 140, App
37	21	91.3	9	4	US-09-919-048-144	Sequence 144, App
38	21	91.3	10	1	US-08-049-511-10	Sequence 10, Appl
39	21	91.3	10	1	US-08-049-511-11	Sequence 11, Appl
40	21	91.3	10	1	US-08-049-511-13	Sequence 13, Appl
41	21	91.3	10	2	US-08-556-597-92	Sequence 92, Appl
42	21	91.3	10	2	US-08-318-856A-56	Sequence 56, Appl
43	21	91.3	10	3	US-09-043-930-14	Sequence 14, Appl
44	21	91.3	10	4	US-09-258-689-3	Sequence 3, Appli
45	21	91.3	10	4	US-09-195-379-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-07-729-099-12
; Sequence 12, Application US/07729099
; Patent No. 5403581
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Pasamontes, Luis
; TITLE OF INVENTION: Coccidiosis Vaccines
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/729,099
; FILING DATE: 19910712
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Eimeria tenella
; US-07-729-099-12

Query Match 91.3%; Score 21; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||

Db 3 PHG 5

RESULT 2

US-08-257-392-12

; Sequence 12, Application US/08257392

; Patent No. 568513

; GENERAL INFORMATION:

; APPLICANT: Binger, Mary-Helen

; APPLICANT: Pasamontes, Luis

; TITLE OF INVENTION: Coccidiosis Vaccines

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: New Jersey

; COUNTRY: U.S.

; ZIP: 07110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/257,392

; FILING DATE: 09-JUN-1994

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/729,099

; FILING DATE: 12-JUL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Roseman, Catherine R

; REGISTRATION NUMBER: 34,240

; REFERENCE/DOCKET NUMBER: 8514

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-6208

; TELEFAX: (201) 235-3500

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: YES

; FRAGMENT TYPE: internal

; ORIGINAL SOURCE:

; ORGANISM: Eimeria tenella

US-08-257-392-12

Query Match 91.3%; Score 21; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 4.1e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5

Db 3 PHG 5

RESULT 3

US-08-770-035-12

; Sequence 12, Application US/08770035

; Patent No. 6008342

; GENERAL INFORMATION:

; APPLICANT: Binger, Mary-Helen

; APPLICANT: Pasamontes, Luis

; TITLE OF INVENTION: Coccidiosis Vaccines

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann-La Roche Inc.

; STREET: 340 Kingsland Street

; CITY: Nutley

; STATE: New Jersey

; COUNTRY: U.S.

; ZIP: 07110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/770,035

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/729,099

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Roseman, Catherine R

; REGISTRATION NUMBER: 34,240

; REFERENCE/DOCKET NUMBER: 8514

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201) 235-6208

; TELEFAX: (201) 235-3500

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: YES

; FRAGMENT TYPE: internal

; ORIGINAL SOURCE:

; ORGANISM: Eimeria tenella

US-08-770-035-12

Query Match 91.3%; Score 21; DB 3; Length 5;

Best Local Similarity 100.0%; Pred. No. 4.1e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5

Db 3 PHG 5

RESULT 4

US-10-122-246A-64

; Sequence 64, Application US/10122246A

; Patent No. 6767897

; GENERAL INFORMATION:

; APPLICANT: ABAJIAN, Henry B

; APPLICANT: HLAVKA, Joseph J

; APPLICANT: FEIGNER, John P

; TITLE OF INVENTION: NEW THERAPEUTIC USES OF TRI-, TETRA-, PENTA-, AND POLYPEPTIDES

; FILE REFERENCE: D4850-00021

; CURRENT APPLICATION NUMBER: US/10/122,246A

; CURRENT FILING DATE: 2002-04-11

; PRIOR APPLICATION NUMBER: US 09/625,103

; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: US 08/962,962

; PRIOR FILING DATE: 1997-11-04

; PRIOR APPLICATION NUMBER: US 08/432,651

; PRIOR FILING DATE: 1995-05-02

; PRIOR APPLICATION NUMBER: US 08/238,089

; PRIOR FILING DATE: 1994-05-04

; NUMBER OF SEQ ID NOS: 110

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 64

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic Peptide

; FEATURE:

; NAME/KEY: MOD_RES

; LOCATION: (1)..(1)

OTHER INFORMATION: 4-fluorophenylalanine
FEATURE:
NAME/KEY: MOD RES
LOCATION: (2)-(2)
OTHER INFORMATION: 4Hyp
FEATURE:
NAME/KEY: MOD RES
LOCATION: (5)-(5)
OTHER INFORMATION: AMIDATION
US-10-122-246A-64

Query Match 91.3%; Score 21; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 2 PHG 4

RESULT 5
5217869-59
Patent No. 5217869
APPLICANT: KAUVAR, LAWRENCE M.
TITLE OF INVENTION: METHOD TO PRODUCE IMMUNODIAGNOSTIC
REAGENTS
NUMBER OF SEQUENCES: 121
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/255,906
FILING DATE: 11-OCT-1988
SEQ ID NO:59:
LENGTH: 5

Query Match 91.3%; Score 21; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 2 PHG 4

RESULT 6
5217869-59
Patent No. 5217869
APPLICANT: KAUVAR, LAWRENCE M.
TITLE OF INVENTION: METHOD TO PRODUCE IMMUNODIAGNOSTIC
REAGENTS
NUMBER OF SEQUENCES: 121
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/255,906
FILING DATE: 11-OCT-1988
SEQ ID NO:59:
LENGTH: 5

Query Match 91.3%; Score 21; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 2 PHG 4

RESULT 7
US-09-095-407-3
Sequence 3, Application US/09095407
Patent No. 6124137
GENERAL INFORMATION:
APPLICANT: Hutchens, T. William
Yip, Tai-Tung

TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
Detection of Analytes
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,407
FILING DATE: 10-Jun-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,896
FILING DATE: 28-MAY-1993
APPLICATION NUMBER: WO PCT/US94/06064
FILING DATE: 27-MAY-1994
APPLICATION NUMBER: US 08/556,951
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5639-PCT-US-DI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 651-5325
TELEFAX: (713) 651-5246
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-095-407-3

Query Match 91.3%; Score 21; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 4 PHG 6

RESULT 8
US-09-560-915-5
Sequence 5, Application US/09560915
Patent No. 6383764
GENERAL INFORMATION:
APPLICANT: Civelli, Olivier
APPLICANT: Lin, Steven
TITLE OF INVENTION: Therapeutic Compositions and Methods
Relating To Prolactin Releasing Peptide (PrRP)
FILE REFERENCE: P-UC 3534
CURRENT APPLICATION NUMBER: US/09/560,915
CURRENT FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: human GPR10 variant
US-09-560-915-5

```

Query Match          91.3%; Score 21; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      1 PHG 3

RESULT 9
US-09-155-613A-71
; Sequence 71, Application US/091555613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-09-155-613A-71

Query Match          91.3%; Score 21; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      2 PHG 4

RESULT 10
PCT-US95-04567-11
; Sequence 11, Application PC/TUS9504567
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04567
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: US 08/227,357
; FILING DATE: 13-APR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UOAB025P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04567-11

Query Match          91.3%; Score 21; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      1 PHG 3

RESULT 11
US-08-776-265-12
; Sequence 12, Application US/08776265
; Patent No. 6001631
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

US-08-776-265-12

Query Match 91.3%; Score 21; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 3 PHG 5

RESULT 12

US-09-398-184-12
; Sequence 12, Application US/09398184
; Patent No. 6649394

; GENERAL INFORMATION:

; APPLICANT: BLANCHE, Francis
; CAMERON, Beatrice
; CROUZET, Joel
; FAMECHON, Alain
; FERRERO, Lucia

; TITLE OF INVENTION: No. 6649394el Topoisomerase IV, Corresponding
; Nucleotide Sequences and Uses Thereof

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.

; STREET: 1300 I. Street, N.W., Suite 700

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/398,184

; FILING DATE: 17-Sep-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/776,265

; FILING DATE: 24-JAN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Einaudi, Carol P.

; REGISTRATION NUMBER: 32,220

; REFERENCE/DOCKET NUMBER: 03806.0394-00000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4444

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-398-184-12

Query Match

Best Local Similarity 91.3%; Score 21; DB 4; Length 7;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 3 PHG 5

RESULT 13

US-08-405-647B-13

; Sequence 13, Application US/08405647B

; Patent No. 6124262

; GENERAL INFORMATION:

; APPLICANT: Sherman, Irwin W.

; APPLICANT: Crandall, Ian E.

; APPLICANT: Sholet, Stephen B.

; APPLICANT: Thevenin, Bernard Jean-Marie

; TITLE OF INVENTION: Compositions and Methods for Reducing

; Adhesiveness of Defective Red Blood Cells

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/405,647B

; FILING DATE: 17-MAR-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Weber, Kenneth A.

; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 02307E-068700US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-405-647B-13

Query Match 91.3%; Score 21; DB 3; Length 8;

Best Local Similarity 100.0%; Pred. No. 4.1e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 6 PHG 8

RESULT 14

US-08-985-499-13

; Sequence 13, Application US/08985499

; Patent No. 6191103

; GENERAL INFORMATION:

; APPLICANT: Sholet, Stephen B.

; APPLICANT: Sherman, Irwin

; APPLICANT: von Andrian, Ulrich

; TITLE OF INVENTION: Methods for Enhancing Thrombolysis in a

; Mammal

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,499
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Laurence J.
; REGISTRATION NUMBER: 35,551
; REFERENCE/DOCKET NUMBER: 02307E-084500US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-985-499-13

Query Match 91.3%; Score 21; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 6 PHG 8

RESULT 15
US-09-125-099-2
; Sequence 2, Application US/09125099A
; Patent No. 6207150
; GENERAL INFORMATION:
; APPLICANT: CROUZET, Joel
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUDER, Michel
; APPLICANT: CAMERON, Beatrice
; TITLE OF INVENTION: VARIANTS OF THYMIDINE KINASE, RELATED NUCLEIC ACIDS
; TITLE OF INVENTION: SEQUENCES AND THEIR USE IN GENIC THERAPY
; FILE REFERENCE: ST96010-US
; CURRENT APPLICATION NUMBER: US/09/125,099A
; CURRENT FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: PCT/FR97/00193
; EARLIER FILING DATE: 1997-01-31
; EARLIER APPLICATION NUMBER: FR96/01603
; EARLIER FILING DATE: 1996-02-09
; EARLIER APPLICATION NUMBER: FR96/09709
; EARLIER FILING DATE: 1996-08-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8
; TYPE: PRT
; ORGANISM: herpes simplex virus 7
;
US-09-125-099-2

Query Match 91.3%; Score 21; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 2 PHG 4

RESULT 16
US-09-239-043D-1149
; Sequence 1149, Application US/09239043D
; Patent No. 6689363
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John

```

```

; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1149
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
;
US-09-239-043D-1149

Query Match 91.3%; Score 21; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 3 PHG 5

RESULT 17
US-09-239-043D-1209
; Sequence 1209, Application US/09239043D
; Patent No. 6689363
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10

```



```

; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1209
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1209

Query Match          91.3%; Score 21; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      1 PHG 3

RESULT 18
US-09-112-956-1
; Sequence 1, Application US/09112956
; Patent No. 6750025
; GENERAL INFORMATION:
; APPLICANT: Hammond, David
; APPLICANT: Medina, Emma
; TITLE OF INVENTION: A NOVEL METHOD OF DETECTING AND ISOLATING PRION
; TITLE OF INVENTION: PROTEIN AND VARIANTS THEREOF, AND NOVEL METHODS OF
; TITLE OF INVENTION: DIAGNOSING AND TREATING PRION DISEASES
; FILE REFERENCE: 92053/26
; CURRENT APPLICATION NUMBER: US/09/112,956
; CURRENT FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-112-956-1

Query Match          91.3%; Score 21; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      5 PHG 7

RESULT 19
PCT-US96-03180-13
; Sequence 13, Application PC/TUS9603180
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University
; APPLICANT: of California
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REDUCING
```

```

; TITLE OF INVENTION: ADHESIVENESS OF DEFECTIVE RED BLOOD CELLS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Robbins, Berliner & Carson
; STREET: 201 North Figueroa Street, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90012-2628
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03180
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berliner, Robert
; REGISTRATION NUMBER: 20,121
; REFERENCE/DOCKET NUMBER: 5555-370
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 977-1001
; TELEFAX: (213) 977-1003
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US96-03180-13

Query Match          91.3%; Score 21; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      6 PHG 8

RESULT 20
US-09-510-738A-102
; Sequence 102, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 102
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-510-738A-102

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      6 PHG 8
```

```
RESULT 21
US-09-510-738A-124
; Sequence 124, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 124
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-510-738A-124

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      6 PHG 8

RESULT 22
US-09-510-738A-125
; Sequence 125, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-510-738A-125

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      6 PHG 8

RESULT 23
US-09-510-738A-140
; Sequence 140, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 140
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-510-738A-140

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      4 PHG 6

RESULT 24
US-09-510-738A-144
; Sequence 144, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-510-738A-144

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      6 PHG 8

RESULT 25
US-09-861-966-102
; Sequence 102, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 102
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-861-966-102
```

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 6 PHG 8

RESULT 26
US-09-861-966-124
; Sequence 124, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 124
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-861-966-124

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 6 PHG 8

RESULT 27
US-09-861-966-125
; Sequence 125, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-861-966-125

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 4 PHG 6

RESULT 28
US-09-861-966-140

; Sequence 140, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 140
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-861-966-140

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 6 PHG 8

RESULT 29
US-09-861-966-144
; Sequence 144, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-861-966-144

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 4 PHG 6

RESULT 30
US-09-239-043D-1128
; Sequence 1128, Application US/09239043D
; Patent No. 6689363
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.

```
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1128
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1128

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          3 PHG 5
           |||
Db          2 PHG 4

RESULT 31
US-09-239-043D-1667
; Sequence 1667, Application US/09239043D
; Patent No. 6689363
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0060007
; CURRENT APPLICATION NUMBER: US/09/239,043D
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 09/189,702
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: US 08/978,291
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: US 08/820,360
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: US 60/013,363
; PRIOR FILING DATE: 1996-03-13
```

```
; PRIOR APPLICATION NUMBER: US 08/461,603
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: US 08/344,824
; PRIOR FILING DATE: 1994-11-23
; PRIOR APPLICATION NUMBER: US 08/278,634
; PRIOR FILING DATE: 1994-07-21
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/197,484
; PRIOR FILING DATE: 1994-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2579
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1667
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1667

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          3 PHG 5
           |||
Db          2 PHG 4

RESULT 32
US-09-870-089B-13
; Sequence 13, Application US/09870089B
; Patent No. 6737062
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER
; FILE REFERENCE: 68126881209900
; CURRENT APPLICATION NUMBER: US/09/870,089B
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ATF4/CREB-2
US-09-870-089B-13

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          3 PHG 5
           |||
Db          5 PHG 7

RESULT 33
US-09-919-048-102
; Sequence 102, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
```

```

; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 102
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-919-048-102

Query Match      91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      4 PHG 6

RESULT 34
US-09-919-048-124
; Sequence 124, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 124
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-919-048-124

Query Match      91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      6 PHG 8

RESULT 35
US-09-919-048-125
; Sequence 125, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-919-048-125

Query Match      91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      6 PHG 8

RESULT 36
US-09-919-048-140
; Sequence 140, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 140
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-919-048-140

Query Match      91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      4 PHG 6

RESULT 37
US-09-919-048-144
; Sequence 144, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-919-048-144

Query Match      91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      4 PHG 6
```

RESULT 38
US-08-049-511-10
; Sequence 10, Application US/08049511
; Patent No. 5316775
; GENERAL INFORMATION:
; APPLICANT: WAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H
; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAWISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING
; TITLE OF INVENTION: HEPATITIS B INFECTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,511
; FILING DATE: 19930419
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/804,844
; FILING DATE: 04-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,267
; FILING DATE: 11-JUL-1991
; APPLICATION NUMBER: US/07/228,364
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-049-511-10

Query Match 91.3%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 8 PHG 10

RESULT 39
US-08-049-511-11
; Sequence 11, Application US/08049511
; Patent No. 5316775
; GENERAL INFORMATION:
; APPLICANT: WAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H

; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAWISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING
; TITLE OF INVENTION: HEPATITIS B INFECTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,511
; FILING DATE: 19930419
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/804,844
; FILING DATE: 04-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,267
; FILING DATE: 11-JUL-1991
; APPLICATION NUMBER: US/07/228,364
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-049-511-11

Query Match 91.3%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 8 PHG 10

RESULT 40
US-08-049-511-13
; Sequence 13, Application US/08049511
; Patent No. 5316775
; GENERAL INFORMATION:
; APPLICANT: WAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H
; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAWISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING
; TITLE OF INVENTION: HEPATITIS B INFECTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960

; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,511
; FILING DATE: 19930419
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/804,844
; FILING DATE: 04-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,267
; FILING DATE: 11-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/228,364
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-049-511-13

Query Match 91.3%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5

|||

Db 8 PHG 10

Search completed: June 15, 2005, 14:24:18
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-7
Perfect score: 23
Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	91.3	4	17 US-10-476-861A-31	Sequence 31, Appl
2	21	91.3	4	17 US-10-923-940-25	Sequence 25, Appl
3	21	91.3	5	9 US-09-817-661-32	Sequence 32, Appl
4	21	91.3	5	14 US-10-074-225A-7	Sequence 7, Appli
5	21	91.3	5	14 US-10-074-225A-8	Sequence 8, Appli
6	21	91.3	5	14 US-10-074-225A-9	Sequence 9, Appli
7	21	91.3	5	14 US-10-074-225A-10	Sequence 10, Appl
8	21	91.3	5	14 US-10-074-225A-11	Sequence 11, Appl
9	21	91.3	6	9 US-09-932-161-5	Sequence 5, Appli
10	21	91.3	6	13 US-10-156-820-71	Sequence 71, Appl
11	21	91.3	6	13 US-10-026-237-25	Sequence 25, Appl

12	21	91.3	6	14 US-10-153-312A-1	Sequence 1, Appli
13	21	91.3	6	14 US-10-096-777-5	Sequence 5, Appli
14	21	91.3	6	16 US-10-699-088-232	Sequence 232, App
15	21	91.3	6	16 US-10-699-088-240	Sequence 240, App
16	21	91.3	6	16 US-10-699-088-758	Sequence 758, App
17	21	91.3	6	16 US-10-699-088-765	Sequence 765, App
18	21	91.3	6	16 US-10-699-088-957	Sequence 957, App
19	21	91.3	6	16 US-10-699-088-964	Sequence 964, App
20	21	91.3	6	16 US-10-699-113-46	Sequence 46, Appl
21	21	91.3	6	16 US-10-699-113-53	Sequence 53, Appl
22	21	91.3	6	16 US-10-699-113-232	Sequence 232, App
23	21	91.3	6	16 US-10-699-113-240	Sequence 240, App
24	21	91.3	6	16 US-10-699-113-758	Sequence 758, App
25	21	91.3	6	16 US-10-699-113-765	Sequence 765, App
26	21	91.3	6	17 US-10-699-114-232	Sequence 232, App
27	21	91.3	6	17 US-10-699-114-240	Sequence 240, App
28	21	91.3	6	17 US-10-699-114-758	Sequence 758, App
29	21	91.3	6	17 US-10-699-114-765	Sequence 765, App
30	21	91.3	6	17 US-10-699-114-957	Sequence 957, App
31	21	91.3	6	17 US-10-699-114-964	Sequence 964, App
32	21	91.3	6	17 US-10-806-924-9	Sequence 9, Appli
33	21	91.3	6	17 US-10-806-924-16	Sequence 16, Appl
34	21	91.3	6	17 US-10-806-924-195	Sequence 195, App
35	21	91.3	6	17 US-10-806-924-203	Sequence 203, App
36	21	91.3	6	17 US-10-806-924-721	Sequence 721, App
37	21	91.3	6	17 US-10-806-924-728	Sequence 728, App
38	21	91.3	7	10 US-09-954-385-70	Sequence 70, Appl
39	21	91.3	7	10 US-09-954-385-377	Sequence 377, App
40	21	91.3	7	15 US-10-646-807-12	Sequence 12, Appl
41	21	91.3	7	17 US-10-912-512-70	Sequence 70, Appl
42	21	91.3	7	17 US-10-912-512-377	Sequence 377, App
43	21	91.3	7	17 US-10-235-043-70	Sequence 70, Appl
44	21	91.3	7	17 US-10-235-043-377	Sequence 377, App
45	21	91.3	8	9 US-09-734-357-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-476-861A-31
; Sequence 31, Application US/10476861A
; Publication No. US20050084902A1
; GENERAL INFORMATION:
; APPLICANT: NOKIHARA , Kiyoshi et al.
; TITLE OF INVENTION: PEPTIDE-IMMOBILIZED SUBSTRATE AND METHOD FOR MEASURING TARGET
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 0760-0325P
; CURRENT APPLICATION NUMBER: US/10/476,861A
; CURRENT FILING DATE: 2003-11-06
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide fragment of SEQ ID
; OTHER INFORMATION: NO: 7 and example 3
US-10-476-861A-31

Query Match 91.3%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 2 PHG 4

RESULT 2
US-10-923-940-25
; Sequence 25, Application US/10923940

```

; Publication No. US20050113297A1
; GENERAL INFORMATION:
; APPLICANT: Francois, et al.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Phagocytosis or Phagocyte
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 2005284-0010
; CURRENT APPLICATION NUMBER: US/10/923,940
; CURRENT FILING DATE: 2004-08-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Identified by phage display. Not from any organism
US-10-923-940-25

Query Match          91.3%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 3
US-09-817-661-32
; Sequence 32, Application US/09817661
; Patent No. US20020076692A1
; GENERAL INFORMATION:
; APPLICANT: Osbourn, Jane
; APPLICANT: Holet, Thor
; TITLE OF INVENTION: Improvements to ribosome display
; FILE REFERENCE: 84633
; CURRENT APPLICATION NUMBER: US/09/817,661
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/193,802
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutagenized
; OTHER INFORMATION: sequence
US-09-817-661-32

Query Match          91.3%; Score 21; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 4
US-10-074-225A-7
; Sequence 7, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A

Query Match          91.3%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 5
US-10-074-225A-8
; Sequence 8, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-8

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 6
US-10-074-225A-9
; Sequence 9, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
```

```

; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(2)
; OTHER INFORMATION: Xaa at positions 1 and 2 can be either His or Pro
US-10-074-225A-7

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 5
US-10-074-225A-8
; Sequence 8, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-8

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 6
US-10-074-225A-9
; Sequence 9, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
```

```

; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-9

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      3 PHG 5

RESULT 7
US-10-074-225A-10
; Sequence 10, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-10

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      3 PHG 5

RESULT 8
US-10-074-225A-11
; Sequence 11, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370

```

```

; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-11

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      3 PHG 5

RESULT 9
US-09-932-161-5
; Sequence 5, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human GPR10 variant
US-09-932-161-5

Query Match          91.3%; Score 21; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      1 PHG 3

RESULT 10
US-10-156-820-71
; Sequence 71, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0

```

```

; SEQ ID NO 71
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-10-156-820-71

Query Match          91.3%; Score 21; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 11
US-10-026-237-25
; Sequence 25, Application US/10026237
; Publication No. US20020173452A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Cytostatic-Glycoconjugates having specifically cleavable linking
; FILE REFERENCE: Lea 34 491
; CURRENT APPLICATION NUMBER: US/10/026,237
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: EP 00128402.5
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Amino Acid sequence not obtained from organism, but synthesized m
; OTHER INFORMATION: annually
US-10-026-237-25

Query Match          91.3%; Score 21; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 12
US-10-153-312A-1
; Sequence 1, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
US-10-153-312A-1
```

```

Query Match          91.3%; Score 21; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      4 PHG 6

RESULT 13
US-10-096-777-5
; Sequence 5, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human GPR10 variant
US-10-096-777-5

Query Match          91.3%; Score 21; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 14
US-10-699-088-232
; Sequence 232, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-232

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Qy 3 PHG 5
|||
Db 1 PHG 3

RESULT 15
US-10-699-088-240
; Sequence 240, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 240
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-240

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 1 PHG 3

RESULT 16
US-10-699-088-758
; Sequence 758, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 758
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-758

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 2 PHG 4

RESULT 17
US-10-699-088-765
; Sequence 765, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-765

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 3 PHG 5

RESULT 18
US-10-699-088-957
; Sequence 957, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 957
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-957

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 2 PHG 4

RESULT 19

```
US-10-699-088-964
; Sequence 964, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 964
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-964

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      3 PHG 5

RESULT 20
US-10-699-113-46
; Sequence 46, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-46

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      2 PHG 4

RESULT 21
US-10-699-113-53
; Sequence 53, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-53

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      3 PHG 5

RESULT 22
US-10-699-113-232
; Sequence 232, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-232

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
Db      1 PHG 3

RESULT 23
US-10-699-113-240
; Sequence 240, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
```

; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 240
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-240

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 24
US-10-699-113-758
; Sequence 758, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 758
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-758

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 2 PHG 4

RESULT 25
US-10-699-113-765
; Sequence 765, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10

; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-765

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 3 PHG 5

RESULT 26
US-10-699-114-232
; Sequence 232, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METHC
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-232

Query Match 91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 27
US-10-699-114-240
; Sequence 240, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METHC
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018

```

; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 240
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-240

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 28
US-10-699-114-758
; Sequence 758, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 758
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-758

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 29
US-10-699-114-765
; Sequence 765, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; FILE REFERENCE: 25885-1759

```

```

; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-765

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 30
US-10-699-114-957
; Sequence 957, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METHC
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 957
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-957

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 31
US-10-699-114-964
; Sequence 964, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde

```

```
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 964
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-964

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      3 PHG 5

RESULT 32
US-10-806-924-9
; Sequence 9, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-9

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      3 PHG 5

RESULT 33
US-10-806-924-16
; Sequence 16, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0

; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 964
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-964

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      3 PHG 5

RESULT 34
US-10-806-924-195
; Sequence 195, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-195

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 PHG 5
      |||
Db      1 PHG 3

RESULT 35
US-10-806-924-203
; Sequence 203, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-203

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 3 PHG 5
|||
Db 1 PHG 3

RESULT 36
US-10-806-924-721
; Sequence 721, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 721
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-721

Query Match 91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 2 PHG 4

RESULT 37
US-10-806-924-728
; Sequence 728, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 728
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-728

Query Match 91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 3 PHG 5

RESULT 38
US-09-954-385-70
; Sequence 70, Application US/09954385
; Publication No. US20030100467A1
; GENERAL INFORMATION:
; APPLICANT: Aehle, Wolfgang
; APPLICANT: Baldwin, Toby L.
; APPLICANT: Van Gastel, Franciscus J.C.
; APPLICANT: Janssen, Giselle G.
; APPLICANT: Murray, Christopher J.
; APPLICANT: Wang, Huaming
; APPLICANT: Winetzky, Deborah S.
; TITLE OF INVENTION: Binding Phenol Oxidizing Enzyme-peptide
; TITLE OF INVENTION: Complexes
; FILE REFERENCE: GC690
; CURRENT APPLICATION NUMBER: US/09/954,385
; CURRENT FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 433
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: binding peptide
US-09-954-385-70

Query Match 91.3%; Score 21; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 5 PHG 7

RESULT 40
US-10-646-807-12
; Sequence 12, Application US/10646807
; Publication No. US20040077005A1
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis

CAMERON, Beatrice
 CROUZET, Joel
 FAMECHON, Alain
 FERRERO, Lucia
 TITLE OF INVENTION: Novel Topoisomerase IV, Corresponding
 Nucleotide Sequences and Uses Thereof
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 Dunner, L.L.P.
 STREET: 1300 I. Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/646,807
 FILING DATE: 25-Aug-2003
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/776,265
 FILING DATE: 24-JAN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Einaudi, Carol P.
 REGISTRATION NUMBER: 32,220
 REFERENCE/DOCKET NUMBER: 03806.0394-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4444
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-646-807-12

Query Match 91.3%; Score 21; DB 15; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.5e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 Db 3 PHG 5

Search completed: June 15, 2005, 14:50:17
 Job time : 80.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-7
Perfect score: 23
Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	21	91.3	18	2 I78841	thrombopoietin rec
2	21	91.3	23	2 A44626	dihydropyrimidine
3	21	91.3	27	2 A27108	larval-specific li
4	21	91.3	32	2 S11610	ribosomal protein
5	21	91.3	34	2 C31514	hemopexin - chicke
6	21	91.3	36	2 E41080	rbpL protein - Rho
7	21	91.3	38	2 G75398	conserved hypothet
8	21	91.3	40	2 S12207	hypothetical prote
9	21	91.3	41	2 A71278	hypothetical prote
10	21	91.3	42	2 T18887	hypothetical prote
11	21	91.3	44	2 H81370	50S ribosomal prot
12	21	91.3	46	2 A99802	hypothetical prote
13	21	91.3	46	2 E97985	hypothetical solub
14	21	91.3	48	2 T50885	hypothetical prote
15	21	91.3	52	2 H81124	protein kinase AK1
16	21	91.3	54	2 S66330	hypothetical prote
17	21	91.3	56	2 B86823	hypothetical prote
18	21	91.3	57	2 S57791	probable beta-gluc
19	21	91.3	57	2 T49863	hypothetical prote
20	21	91.3	57	2 A33811	tracheobronchial m
21	21	91.3	57	2 C97837	(p)ppGpp 3-pyropho
22	21	91.3	58	2 A45765	cre protein - Stre
23	21	91.3	59	2 T03588	pyrophosphate-fruc
24	21	91.3	59	2 A86485	protein F28J9.16 [
25	21	91.3	59	2 D82459	hypothetical prote
26	21	91.3	60	2 A84213	hypothetical prote
27	21	91.3	60	2 G82472	hypothetical prote
28	21	91.3	61	1 ERAD64	early E3 6.4K prot
29	21	91.3	64	2 E86800	prophage pi3 prote

30	21	91.3	64	2 PQ0277	hypothetical prote
31	21	91.3	65	2 S10211	hypothetical prote
32	21	91.3	65	2 G75258	conserved hypothet
33	21	91.3	65	2 S39675	ywbE protein - Bac
34	21	91.3	67	2 B31048	phosphotransferase
35	21	91.3	67	2 T25844	hypothetical prote
36	21	91.3	67	2 AC1037	probable phage tai
37	21	91.3	67	2 AG0927	probable phage tai
38	21	91.3	67	2 B55855	tail protein - pha
39	21	91.3	67	2 T17586	hypothetical prote
40	21	91.3	68	2 G95330	hypothetical prote
41	21	91.3	68	2 AH1989	hypothetical prote
42	21	91.3	69	2 AG2520	hypothetical prote
43	21	91.3	70	2 B72286	ferredoxin - Therm
44	21	91.3	70	2 T32842	hypothetical prote
45	21	91.3	71	2 C75627	hypothetical prote

ALIGNMENTS

RESULT 1
I78841
thrombopoietin receptor - mouse (fragment)
C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I78841
R;Alexander, W.S.; Dunn, A.R.
Oncogene 10, 795-803, 1995
A;Title: Structure and transcription of the genomic locus encoding murine c-Mpl, a recept
A;Reference number: I58350; MUID:95166571; PMID:7862460
A;Accession: I78841
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-18 <RES>
A;Cross-references: GB:S76842; NID:g912990; PIDN:AAB33462.1; PID:g912991
C;Genetics:
A;Gene: c-mpl1

Query Match 91.3%; Score 21; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 2 PHG 4

RESULT 2
A44626
dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - bovine (fragment)
N;Alternate names: dihydrothymine dehydrogenase; dihydouracil dehydrogenase (NADP+)
C;Species: Bos primigenius taurus (cattle)
C;Date: 02-Aug-1994 #sequence_revision 23-Mar-1995 #text_change 03-Jun-2002
C;Accession: A44626; A41066
R;Porter, D.J.T.; Chestnut, W.G.; Merrill, B.M.; Spector, T.
J. Biol. Chem. 267, 5236-5242, 1992
A;Title: Mechanism-based inactivation of dihydropyrimidine dehydrogenase by 5-ethynylurac
A;Reference number: A44626; MUID:92184771; PMID:1544906
A;Accession: A44626
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-23 <POR>
R;Porter, D.J.T.; Chestnut, W.G.; Taylor, L.C.E.; Merrill, B.M.; Spector, T.
J. Biol. Chem. 266, 19988-19994, 1991
A;Title: Inactivation of dihydropyrimidine dehydrogenase by 5-iodouracil.
A;Reference number: A41066; MUID:92041818; PMID:1939061
A;Accession: A41066
A;Molecule type: protein
A;Residues: 13-23 <PO2>
C;Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hon
C;Keywords: NADP; oxidoreductase
F;16/Binding site: substrate (Cys) #status experimental

Query Match 91.3%; Score 21; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 17 PHG 19

RESULT 3
A27108
larval-specific lipoprotein - honeybee (fragment)
C;Species: Apis mellifera (honeybee)
C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 09-Jul-2004
C;Accession: A27108
R;Shipman, B.A.; Ryan, R.O.; Schmidt, J.O.; Law, J.H.
Biochemistry 26, 1885-1889, 1987
A;Title: Purification and properties of a very high density lipoprotein from the hemolymph of honeybees
A;Reference number: A27108; MUID:87242376; PMID:3109474
A;Accession: A27108
A;Molecule type: protein
A;Residues: 1-27 <SHI>
A;Cross-references: UNIPROT:P09355
C;Keywords: lipid binding; lipoprotein

Query Match 91.3%; Score 21; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 5 PHG 7

RESULT 4
S11610
ribosomal protein S4.eR [validated] - Halobacterium salinarum (fragment)
N;Alternate names: ribosomal protein HS5
C;Species: Halobacterium salinarum
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: S11610
R;Yaguchi, M.; Visentin, L.P.; Zuker, M.; Matheson, A.T.; Roy, C.; Strom, A.R.
Zbl. Bakt. Hyg. I. Abt. Orig. C 3, 200-208, 1982
A;Title: Amino-terminal sequences of ribosomal proteins from the 30S subunit of archaea
A;Reference number: S11609
A;Accession: S11610
A;Molecule type: protein
A;Residues: 1-32 <YAG>
A;Cross-references: UNIPROT:Q7M553
A;Note: the protein is designated as ribosomal protein HS5
A;Note: the source is designated as Halobacterium cutirubrum
C;Superfamily: rat ribosomal protein S4
C;Keywords: protein biosynthesis; ribosome

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 30 PHG 32

RESULT 5
C31514
hemopexin - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: C31514
R;Wellner, D.; Cheng, K.C.; Muller-Eberhard, U.
Biochem. Biophys. Res. Commun. 155, 622-625, 1988
A;Title: N-terminal amino acid sequences of the hemopexins from chicken, rat and rabbit.

A;Reference number: A90148; MUID:88339942; PMID:3421961
A;Accession: C31514
A;Molecule type: protein
A;Residues: 1-34 <WE3>
A;Cross-references: UNIPROT:P20057

Query Match 91.3%; Score 21; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 16 PHG 18

RESULT 6
E41080
rbpL protein - Rhodobacter sphaeroides (fragment)
C;Species: Rhodobacter sphaeroides
C;Date: 12-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 09-Jul-2004
C;Accession: E41080
R;Chen, J.H.; Gibson, J.L.; McCue, L.A.; Tabita, F.R.
J. Biol. Chem. 266, 20447-20452, 1991
A;Title: Identification, expression, and deduced primary structure of transketolase and transketolase-related proteins from Rhodospirillum rubrum
A;Reference number: A41080; MUID:92041881; PMID:1939098
A;Accession: E41080
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-36 <CHE>
A;Cross-references: UNIPROT:P29278; GB:M68914; NID:g151988; PIDN:AAA26158.1; PID:g551958

Query Match 91.3%; Score 21; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 25 PHG 27

RESULT 7
G75398
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75398
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75398
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-38 <WHI>
A;Cross-references: UNIPROT:Q9RUH4; GB:AE001986; GB:AE000513; NID:g6459162; PIDN:AAF1097
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1412
A;Map position: 1

Query Match 91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 31 PHG 33

RESULT 8
S12207

hypothetical protein (B2 element) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C:Accession: S12207

R;Oberbaeumer, I.
submitted to MIPS, January 1991
A:Reference number: S12205
A:Accession: S12207
A:Molecule type: mRNA
A:Residues: 1-40 <OBE>
A:Cross-references: UNIPROT:Q8K081; UNIPROT:Q8K0H2; UNIPROT:Q8RIU0; UNIPROT:Q99JK9; EMBL

Query Match 91.3%; Score 21; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 22 PHG 24

RESULT 9
A71278
hypothetical protein TP0811 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: A71278

R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: A71278
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-41 <COL>

A:Cross-references: UNIPROT:O83788; GB:AE001252; GB:AE000520; NID:g3323119; PIDN:AAC6578
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0811

Query Match 91.3%; Score 21; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 36 PHG 38

RESULT 10
T18887
hypothetical protein C03D6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18887

R;Burton, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19038
A:Accession: T18887
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-42 <WIL>
A:Cross-references: UNIPROT:Q17601; EMBL:Z75525; PIDN:CAA99767.1; GSPDB:GN00019; CESP:CO
A:Experimental source: clone C03D6
C:Genetics:
A:Gene: CESP:C03D6.2
A:Map position: 1

Query Match 91.3%; Score 21; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 14 PHG 16

RESULT 11
H81370
50S ribosomal protein L34 Cj0961c [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81370
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-44 <PAR>
A:Cross-references: UNIPROT:Q9PNX4; GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73218
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: rpmH; Cj0961c
C:Superfamily: Escherichia coli ribosomal protein L34

Query Match 91.3%; Score 21; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 7 PHG 9

RESULT 12
A99802
hypothetical protein ECS1385 [imported] - Escherichia coli (strain O157:H7, substrain RIN
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A99802
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A99802
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-46 <HAY>
A:Cross-references: UNIPROT:Q8X2G6; GB:BA000007; PIDN:BA34808.1; PID:gl3360845; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS1385

Query Match 91.3%; Score 21; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 11 PHG 13

RESULT 13
E97985
hypothetical protein spr0909 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: E97985
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Be
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc

Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: E97985
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <KUR>
A;Cross-references: UNIPROT:Q8CYU7; GB:AE007317; PIDN:AAK99713.1; PID:g15458516; GSPDB:G
C;Genetics:
A;Gene: spr0909

Query Match 91.3%; Score 21; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 23 PHG 25

RESULT 14
T50885
hypothetical soluble protein [imported] - Rubrivivax gelatinosus
C;Species: Rubrivivax gelatinosus
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C;Accession: T50885
R;Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
submitted to the EMBL Data Library, November 1999
A;Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn
A;Reference number: Z25270
A;Accession: T50885
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-48 <NAG>
A;Cross-references: UNIPROT:Q9JPB8; EMBL:AB034704; PIDN:BAA94038.1
A;Experimental source: strain IL144
C;Genetics:
A;Note: ORF48

Query Match 91.3%; Score 21; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 46 PHG 48

RESULT 15
H81124
hypothetical protein NMB1079 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: H81124
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: H81124
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-52 <TET>
A;Cross-references: UNIPROT:Q9JZF4; GB:AE002458; GB:AE002098; NID:g7226311; PIDN:AAF4147
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1079

Query Match 91.3%; Score 21; DB 2; Length 52;

Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 13 PHG 15

RESULT 16
S66330
protein kinase AK17 (EC 2.7.1.-) - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004
C;Accession: S66330; S58271
R;Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
Plant Mol. Biol. 29, 551-565, 1995
A;Title: Differential accumulation of the transcripts of 22 novel protein kinase genes ir
A;Reference number: S66314; MUID:96123233; PMID:8534852
A;Accession: S66330
A;Molecule type: DNA
A;Residues: 1-54 <THU>
A;Cross-references: UNIPROT:Q38982; EMBL:X86962; NID:g928899; PIDN:CAA60525.1; PID:g92890
C;Genetics:
A;Gene: AK17
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; protein kinase
F;1-54/Domain: protein kinase homology (fragment) <KIN>

Query Match 91.3%; Score 21; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 49 PHG 51

RESULT 17
B86823
hypothetical protein yqcG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86823
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlic
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ser
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-56 <STO>
A;Cross-references: UNIPROT:Q9CF95; GB:AE005176; PID:g12724591; PIDN:AAK05684.1; GSPDB:GN
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yqcG

Query Match 91.3%; Score 21; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 36 PHG 38

RESULT 18
S57791
probable beta-glucosidase - black mustard (fragment)
C;Species: Brassica nigra (black mustard)
C;Date: 27-Oct-1995 #sequence_revision 10-May-1996 #text_change 09-Jul-2004
C;Accession: S57791
R;Malboobi, M.A.; Lefebvre, D.D.
Plant Mol. Biol. 28, 859-870, 1995

A;Title: Isolation of cDNA clones of genes with altered expression levels in phosphate-
A;Reference number: S57791; MUID:95367647; PMID:7640358
A;Accession: S57791
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-57 <MAL>
A;Cross-references: UNIPROT:O24434
C;Superfamily: Agrobacterium beta-glucosidase

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 30 PHG 32

RESULT 19
T49863
hypothetical protein B24P11.190 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49863
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49863
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <SCH>
A;Cross-references: EMBL:AL356833; GSPDB:GN00116; NCSP:B24P11.190
A;Experimental source: BAC clone B24P11; strain OR74A
C;Genetics:
A;Gene: NCSP:B24P11.190
A;Map position: 6
A;Introns: 9/1

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 35 PHG 37

RESULT 20
A33811
tracheobronchial mucin - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 09-Jul-2004
C;Accession: A33811
R;Rose, M.C.; Kaufman, B.; Martin, B.M.
J. Biol. Chem. 264, 8193-8199, 1989
A;Title: Proteolytic fragmentation and peptide mapping of human carboxyamidomethylated b
A;Reference number: A33811; MUID:89255255; PMID:2656675
A;Accession: A33811
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-57 <ROS>
A;Cross-references: UNIPROT:Q7M4S5

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 11 PHG 13

RESULT 21

C97837
(p)ppGpp 3-pyrophosphohydrolase homolog RC1099 [imported] - Rickettsia conorii (strain M
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: C97837
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: C97837
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <KUR>
A;Cross-references: UNIPROT:Q92GM4; GB:AE006914; PIDN:AAL03637.1; PID:g15620223; GSPDB:G
C;Genetics:
A;Gene: RC1099

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 39 PHG 41

RESULT 22
A45765
cre protein - Streptomyces hygroscopicus
C;Species: Streptomyces hygroscopicus
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 09-Jul-2004
C;Accession: A45765
R;Ogura, M.; Tanaka, T.; Seto, H.; Otake, N.
J. Antibiot. 43, 873-882, 1990
A;Title: Molecular cloning and characterization of the gene conferring curromycin resist
A;Reference number: A45765; MUID:90354306; PMID:2117601
A;Accession: A45765
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-58 <OGU>
A;Cross-references: UNIPROT:P16961; GB:M28599; NID:g153217; PIDN:AAA26721.1; PID:g153218

Query Match 91.3%; Score 21; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 14 PHG 16

RESULT 23
T03588
pyrophosphate-fructose-6-phosphate 1-phosphotransferase homolog - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03588
R;Umeda, M.; Hara, C.; Matsubayashi, Y.; Li, H.; Lui, Q.; Tadokoro, F.; Aotsuka, S.; Uchi
Plant Mol. Biol. 25, 469-478, 1994
A;Title: Expressed sequence tags from cultured cells of rice (Oryza sativa L.) under stre
A;Reference number: Z14962; MUID:94325471; PMID:8049371
A;Accession: T03588
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-59 <UME>
A;Cross-references: UNIPROT:Q42988; EMBL:D17765; NID:g455496; PIDN:BAA04611.1; PID:g87191
A;Experimental source: callus
C;Genetics:
A;Note: AK99
C;Superfamily: pyrophosphate-dependent phosphofructokinase, PfpB type; 6-phosphofructokir

Query Match 91.3%; Score 21; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||

Db 26 PHG 28

RESULT 24
A86485

protein F28J9.16 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86485

R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maity, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86485

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-59 <STO>

A;Cross-references: UNIPROT:Q9S9R3; GB:AE005172; NID:g6272386; PIDN:AAF06092.1; GSPDB:GN

A;Gene: F28J9.16

A;Map position: 1

Query Match 91.3%; Score 21; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||

Db 12 PHG 14

RESULT 25
D82459

hypothetical protein VCA0435 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: D82459

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82459

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-59 <HEI>

A;Cross-references: UNIPROT:Q9KMC7; GB:AE004377; GB:AE003853; NID:g9657831; PIDN:AAF9634

A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor

C;Genetics:

A;Gene: VCA0435

A;Map position: 2

Query Match 91.3%; Score 21; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||

Db 5 PHG 7

RESULT 26
A84213

hypothetical protein Vng0548c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84213

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li;
A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: A84213

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-60 <STO>

A;Cross-references: UNIPROT:Q9HRT9; GB:AE004437; NID:gi0580147; PIDN:AAG19069.1; GSPDB:GN

C;Genetics:

A;Gene: VNG0548C

Query Match 91.3%; Score 21; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||

Db 46 PHG 48

RESULT 27
G82472

hypothetical protein VCA0336 VCA0353 VCA0297 [imported] - Vibrio cholerae (strain N16961
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82472; A82471; A82477

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: G82472

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-60 <HEI>

A;Cross-references: UNIPROT:O31021; GB:AE004372; GB:AE003853; NID:g9657741; PIDN:AAF9624

A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor

A;Accession: A82471

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-60 <HE2>

A;Cross-references: GB:AE004372; GB:AE003853; NID:g9657741; PIDN:AAF96261.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor

A;Accession: A82477

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-60 <HE3>

A;Cross-references: GB:AE004369; GB:AE003853; NID:g9657685; PIDN:AAF96205.1; GSPDB:GN001

A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor

C;Genetics: <GEN1>

A;Gene: VCA0336

A;Map position: 2

C;Genetics: <GEN2>

A;Gene: VCA0353

A;Map position: 2

C;Genetics: <GEN3>

A;Gene: VCA0297

A;Map position: 2

Query Match 91.3%; Score 21; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 6 PHG 8

RESULT 28
ERAD64
early E3 6.4K protein - human adenovirus 35
C;Species: Mastadenovirus h35 (human adenovirus 35)
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: B31162
R;Flomenberg, P.R.; Chen, M.; Horwitz, M.S.
J. Virol. 62, 4431-4437, 1988
A;Title: Sequence and genetic organization of adenovirus type 35 early region 3.
A;Reference number: A93039; MUID:89012230; PMID:3172347
A;Accession: B31162
A;Molecule type: DNA
A;Residues: 1-61 <FLO>
A;Cross-references: UNIPROT:P17591; GB:M23195; NID:g516584; PIDN:AAA42436.1; PID:g516586
C;Superfamily: adenovirus early E3 6.4K protein
C;Keywords: early protein

Query Match 91.3%; Score 21; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 22 PHG 24

RESULT 29
E86800
prophage pi3 protein 32 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: E86800
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s8
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: E86800
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-64 <STO>
A;Cross-references: UNIPROT:Q9CFR0; GB:AE005176; PID:g12724393; PIDN:AAK05503.1; GSPDB:C
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: pi332

Query Match 91.3%; Score 21; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 34 PHG 36

RESULT 30
PQ0277
hypothetical protein 1 - cercopithecine herpesvirus 1 (fragment)
C;Species: cercopithecine herpesvirus 1
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: PQ0277
R;Killeen, A.M.; Harrington, L.; Wall, L.V.M.; Kelly, D.C.
J. Gen. Virol. 73, 195-199, 1992
A;Title: Nucleotide sequence analysis of a homologue of herpes simplex virus type 1 gene
A;Reference number: JQ1406; MUID:92113572; PMID:1309859
A;Accession: PQ0277
A;Molecule type: DNA

A;Residues: 1-64 <KIL>
A;Cross-references: UNIPROT:P30816; GB:S75996; NID:g243000; PIDN:AAB21001.1; PID:g243001
C;Comment: Simian herpes B virus is an alpha herpesvirus which causes a mild, recurring
Query Match 91.3%; Score 21; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 62 PHG 64

RESULT 31
S10211
hypothetical protein, 7.3K - human adenovirus 41
C;Species: Mastadenovirus h41 (human adenovirus 41)
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: S10211
R;Slemenda, S.B.; Pieniazek, N.J.; Velarde Jr., J.; Pieniazek, D.; Luftig, R.B.
Nucleic Acids Res. 18, 3069, 1990
A;Title: Nucleotide sequence of the region coding for 100K and 33K proteins of human ente
A;Reference number: S10206; MUID:90272433; PMID:2349115
A;Accession: S10211
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-65 <SLE>
A;Cross-references: UNIPROT:P23691; EMBL:X52532; NID:g58438; PIDN:CAA36764.1; PID:g58444

Query Match 91.3%; Score 21; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 24 PHG 26

RESULT 32
G75258
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75258
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Lam,
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75258
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-65 <WHI>
A;Cross-references: UNIPROT:Q9RRD1; GB:AB002085; GB:AE000513; NID:g6460383; PIDN:AAF12095
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2561
A;Map position: 1

Query Match 91.3%; Score 21; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 43 PHG 45

RESULT 33
S39675
ywbE protein - Bacillus subtilis
N;Alternate names: hypothetical protein ipa-20r

C;Species: Bacillus subtilis
C;Date: 07-Oct-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S39675; C70051
R;Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. 10, 371-384, 1993
A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region from
A;Reference number: S39655; MUID:95020537; PMID:7934828
A;Accession: S39675
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-65 <GLA>
A;Cross-references: UNIPROT:P39588; EMBL:X73124; NID:g413923; PIDN:CAA51576.1; PID:g4139
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berten
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choc
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: C70051
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-65 <KUN>
A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15861.1; PID:ell186334;
A;Experimental source: strain 168
C;Genetics:
A;Gene: ywbe

Query Match 91.3%; Score 21; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 44 PHG 46

RESULT 34
B31048
phosphotransferase system enzyme II (EC 2.7.1.69), mannitol-specific, factor III - Staph
N;Alternate names: phosphotransferase system enzyme III, mannitol-specific
C;Species: Staphylococcus carnosus
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 09-Jul-2004
C;Accession: B31048
R;Reiche, B.; Frank, R.; Deutscher, J.; Meyer, N.; Hengstenberg, W.
Biochemistry 27, 6512-6516, 1988
A;Title: Staphylococcal phosphoenolpyruvate-dependent phosphotransferase system: purific
carnosus and homology with the enzyme II(mtl) of Escherichia coli.
A;Reference number: A31048; MUID:89118247; PMID:3064811
A;Accession: B31048
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-67 <REI>
A;Cross-references: UNIPROT:PI7876
C;Superfamily: phosphotransferase system mannitol-specific enzyme II factor III; phospho
C;Keywords: phosphoprotein; phosphotransferase
F;35-67/Domain: phosphotransferase system mannitol-specific enzyme II factor III homolog

Query Match 91.3%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 44 PHG 46

RESULT 35
T25844
hypothetical protein M03F4.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T25844
R;Fulton, L.
submitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid M03F4.
A;Reference number: Z20097
A;Accession: T25844
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-67 <FUL>
A;Cross-references: EMBL:U64601; PIDN:AAB04580.1; GSPDB:GN00028; CESP:M03F4.1
A;Experimental source: strain Bristol N2; clone M03F4
C;Genetics:
A;Gene: CESP:M03F4.1
A;Map position: X

Query Match 91.3%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 28 PHG 30

RESULT 36
AC1037
probable phage tail protein STY4622 [imported] - Salmonella enterica subsp. enterica sero
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC1037
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serova
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC1037
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD06744.1; PID:g16505395; GSPDB:GN00176
C;Genetics:
A;Gene: STY4622

Query Match 91.3%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 44 PHG 46

RESULT 37
AG0927
probable phage tail protein [imported] - Salmonella enterica subsp. enterica serovar Typhi
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0927
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD09441.1; PID:gl6504558; GSPDB:GN00176
C;Genetics:
A;Gene: STY3680

Query Match 91.3%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 44 PHG 46

RESULT 38
B55855
tail protein - phage P2
C;Species: phage P2
C;Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C;Accession: B55855
R;Ziermann, R.; Bartlett, B.; Calendar, R.; Christie, G.E.
J. Bacteriol. 176, 4974-4984, 1994
A;Title: Functions involved in bacteriophage P2-induced host cell lysis and identifica
A;Reference number: A55855; MUID:94327465; PMID:8051010
A;Accession: B55855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <ZIE>
A;Cross-references: UNIPROT:P51772; GB:L29304; NID:g508561; PID:g508562
C;Genetics:
A;Gene: X

Query Match 91.3%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 44 PHG 46

RESULT 39
T17586
hypothetical protein a96R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17586
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17586
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-67 <GRA>
A;Cross-references: UNIPROT:Q84417; EMBL:U42580; NID:g4028896; PIDN:AAC96464.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Gene: a96R

Query Match 91.3%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 33 PHG 35

RESULT 40
G95330

hypothetical protein SMa1017 [imported] - Sinorhizobium meliloti (strain 1021) magaplasmi
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: G95330
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: G95330
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-68 <KUR>
A;Cross-references: UNIPROT:Q922D7; GB:AE006469; PIDN:AAK65209.1; PID:g14523656; GSPDB:G
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa1017
A;Genome: plasmid

Query Match 91.3%; Score 21; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 48 PHG 50

Search completed: June 15, 2005, 14:22:43
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-7
Perfect score: 23
Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	91.3	13	2 Q80XX0	Q80xx0 cricetulus
2	21	91.3	16	2 O21922	O21922 streptococc
3	21	91.3	18	2 Q6LD80	Q6ld80 mus sp. thr
4	21	91.3	19	2 Q90RF4	Q90rf4 human immun
5	21	91.3	19	2 Q90RH4	Q90rh4 human immun
6	21	91.3	26	2 Q6R4S9	Q6r4s9 pyrus commu
7	21	91.3	27	1 LSP_APIME	P09355 apis mellif
8	21	91.3	29	1 GALA_ONCMY	P47213 oncorhynchu
9	21	91.3	29	1 SODC_OLEEU	P80740 olea europ
10	21	91.3	29	2 Q6X6Z8	Q6x6z8 enterococcu
11	21	91.3	29	2 Q6X701	Q6x701 enterococcu
12	21	91.3	29	2 Q8R398	Q8r398 mus musculu
13	21	91.3	30	2 Q989K8	Q989k8 rhizobium l
14	21	91.3	30	2 Q9JIC4	Q9jic4 rattus norv
15	21	91.3	31	2 Q9GKL4	Q9gkl4 canis famil
16	21	91.3	31	2 Q72DY1	Q72dy1 desulfovibr
17	21	91.3	32	2 Q7M553	Q7m553 halobacteri
18	21	91.3	32	2 Q6QLS6	Q6qls6 rhodococcus
19	21	91.3	32	2 Q736U2	Q736u2 bacillus ce
20	21	91.3	33	2 O02832	O02832 gorilla gor
21	21	91.3	33	2 Q66384	Q66384 thiobacillu
22	21	91.3	33	2 Q9F709	Q9f709 methylobaci
23	21	91.3	33	2 Q9F710	Q9f710 methylobact
24	21	91.3	33	2 Q9F711	Q9f711 methylobact
25	21	91.3	34	1 HEMO_CHICK	P20057 gallus gall
26	21	91.3	34	2 Q73RE8	Q73re8 treponema d
27	21	91.3	35	2 Q6LBR0	Q6lbr0 mus musculu
28	21	91.3	35	2 Q8CLW4	Q8clw4 yersinia pe
29	21	91.3	36	2 Q8ZWQ4	Q8zwq4 pyrobaculum
30	21	91.3	36	2 Q9URE9	Q9ure9 saccharomyc
31	21	91.3	37	2 Q7TNQ0	Q7tnq0 mus musculu

32	21	91.3	38	2 Q6X700	Q6x700 enterococcu
33	21	91.3	38	2 Q8VJV6	Q8vjv6 mycobacteri
34	21	91.3	38	2 Q81TC9	Q81tc9 bacillus an
35	21	91.3	38	2 Q9RUH4	Q9ruh4 deinococcus
36	21	91.3	39	2 Q6EES4	Q6ees4 homo sapien
37	21	91.3	39	2 Q9NC67	Q9nc67 dictyosteli
38	21	91.3	39	2 Q6X702	Q6x702 enterococcu
39	21	91.3	40	2 Q8ZTN8	Q8ztn8 pyrobaculum
40	21	91.3	40	2 P95768	P95768 staphylococ
41	21	91.3	40	2 Q48916	Q48916 mycobacteri
42	21	91.3	40	2 Q48917	Q48917 mycobacteri
43	21	91.3	40	2 Q48918	Q48918 mycobacteri
44	21	91.3	40	2 Q48951	Q48951 mycobacteri
45	21	91.3	40	2 Q49102	Q49102 mycobacteri

ALIGNMENTS

RESULT 1					
Q80XX0					
ID	Q80XX0	PRELIMINARY;	PRT;	13 AA.	
AC	Q80XX0;				
DT	01-JUN-2003 (TrEMBLrel. 24, Created)				
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE	Hypoxanthine phosphoribosyl transferase (Fragment).				
GN	Name=hprt;				
OS	Cricetulus griseus (Chinese hamster).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;				
OC	Cricetulus.				
OX	NCBI_TaxID=10029;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93024555; PubMed=1383700; DOI=10.1016/0027-5107(92)90198-B;				
RA	Fuscoe J.C.; Zimmerman L.J.; Fekete A.; Setzer R.W.; Rossiter B.J.;				
RT	"Analysis of X-ray-induced HPRT mutations in CHO cells: insertion and deletions.";				
RL	Mutat. Res. 269:171-183(1992).				
DR	EMBL; S46270; AAP13884.1; -.				
DR	GO; GO:0016740; F:transferase activity; IEA.				
KW	Transferase.				
FT	NON TER 13				
SQ	SEQUENCE 13 AA; 1516 MW; 113A46B08648F760 CRC64;				
Query Match 91.3%; Score 21; DB 2; Length 13;					
Best Local Similarity 100.0%; Pred. No. 4e+02;					
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	3 PHG 5				
Db	7 PHG 9				
RESULT 2					
O21922					
ID	O21922	PRELIMINARY;	PRT;	16 AA.	
AC	O21922; O21923;				
DT	01-JAN-1998 (TrEMBLrel. 05, Created)				
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	Integrase (Fragment).				
GN	Name=int;				
OS	Streptococcus thermophilus bacteriophage Sfi21.				
OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.				
OX	NCBI_TaxID=64186;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98008989; PubMed=9344917; DOI=10.1006/viro.1997.8769;				
RA	Bruttin A.; Foley S.; Brussow H.;				
RT	"The site-specific integration system of the temperate Streptococcus thermophilus bacteriophage phisfi21.";				

RL Virology 237:148-158(1997).
DR EMBL; AF013584; AAC48909.1; -.
DR EMBL; AF013587; AAC48910.1; -.
FT NON_TER 1 1
SQ SEQUENCE 16 AA; 1856 MW; 8FA82D3270B9A959 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 14 PHG 16

RESULT 3
Q6LD80
ID Q6LD80 PRELIMINARY; PRT; 18 AA.
AC Q6LD80;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Thrombopoietin receptor (Fragment).
GN Name=c-mpl1;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95166571; PubMed=7862460;
RA Alexander W.S., Dunn A.R.;
RT "Structure and transcription of the genomic locus encoding murine c-
Mpl, a receptor for thrombopoietin.";
RL Oncogene 10:795-803(1995).
DR EMBL; S76842; AAB33462.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1808 MW; E01CB44EF83834E2 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 2 PHG 4

RESULT 4
Q90RF4
ID Q90RF4 PRELIMINARY; PRT; 19 AA.
AC Q90RF4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CG282;
RX MEDLINE=21663216; PubMed=11804559; DOI=10.1089/088922202753394745;
RA Taniguchi Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., Obengui, M'Pandi M., M'Pele P., Harada Y.,
RA Ido E., Hayami M., Ichimura H., Parra H.Joseph.;
RT "Genetic subtypes of HIV type 1 based on the vpu/env sequences in the
Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 18:79-83(2002).
DR EMBL; AF127551; AAK84917.1; -.

FT NON_TER 1 1
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2163 MW; 2AB3630189068236 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 7 PHG 9

RESULT 5
Q90RH4
ID Q90RH4 PRELIMINARY; PRT; 19 AA.
AC Q90RH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96CG17;
RX MEDLINE=21663216; PubMed=11804559; DOI=10.1089/088922202753394745;
RA Taniguchi Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., Obengui, M'Pandi M., M'Pele P., Harada Y.,
RA Ido E., Hayami M., Ichimura H., Parra H.Joseph.;
RT "Genetic subtypes of HIV type 1 based on the vpu/env sequences in the
Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 18:79-83(2002).
DR EMBL; AF127546; AAK84900.1; -.
FT NON_TER 1 1
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2180 MW; 2A81A56D59068236 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 7 PHG 9

RESULT 6
Q6R4S9
ID Q6R4S9 PRELIMINARY; PRT; 26 AA.
AC Q6R4S9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE 1-aminocyclopropene-1-carboxylate synthase 1b (Fragment).
GN Name=DACS1b;
OS Pyrus communis (Pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=23211;
RN [1]
RP SEQUENCE FROM N.A.
RA El Sharkawy I., Lelievre J.M., Latche A., Pech J.C.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY514041; AAS66777.1; -.
DR HSSP; P37821; 1B8G.
FT NON_TER 26 26
SQ SEQUENCE 26 AA; 3117 MW; 2A59C0F5BEA39D13 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 26;

Best Local Similarity 100.0%; Pred. No. 8.2e+02; Mismatches 0; Indels 0; Gaps 0; Matches 3; Conservative 0;

QY 3 PHG 5
Db 12 PHG 14

RESULT 7
LSP_APIME STANDARD; PRT; 27 AA.
AC P09355;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Larval-specific very high density lipoprotein (VHDL) (Fragment).
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=87242376; PubMed=3109474;
RA Shipman B.A., Ryan R.O., Schmidt J.O., Law J.H.;
RT "Purification and properties of a very high density lipoprotein from the hemolymph of the honeybee Apis mellifera.";
RL Biochemistry 26:1885-1889(1987).
CC -!- FUNCTION: Unknown (it might play a role in lipid transport and/or storage protein metabolism during metamorphosis).
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Hemolymph.
CC -!- DEVELOPMENTAL STAGE: Present in high amounts in hemolymph only at the end of larval life.
DR PIR; A27108; A27108.
KW Direct protein sequencing; Hemolymph; Lipoprotein.
FT NON TER 27
SQ SEQUENCE 27 AA; 2885 MW; A9DDF7A75A65D10D CRC64;

Query Match 91.3%; Score 21; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.5e+02; Mismatches 0; Indels 0; Gaps 0; Matches 3; Conservative 0;

QY 3 PHG 5
Db 5 PHG 7

RESULT 8
GALA_ONCMY STANDARD; PRT; 29 AA.
AC P47213;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galanin.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE.
RC TISSUE=Stomach;
RX MEDLINE=95164756; PubMed=7532194;
RA Anglade I., Wang Y., Jensen J., Tramu G., Kah O., Conlon J.M.;
RT "Characterization of trout galanin and its distribution in trout brain and pituitary.";
RL J. Comp. Neurol. 350:63-74(1994).
CC -!- FUNCTION: Contracts smooth muscle of the gastrointestinal and genitourinary tract, regulates growth hormone release, modulates

insulin release, and may be involved in the control of adrenal secretion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the galanin family.
DR InterPro; IPR008174; Galanin.
DR Pfam; PF01296; Galanin; 1.
DR PRINTS; PR00273; GALANIN.
DR ProDom; PD005962; Galanin; 1.
DR PROSITE; PS00861; GALANIN; 1.
KW Amidation; Direct protein sequencing; Hormone; Neuropeptide.
FT MOD RES 29
SQ SEQUENCE 29 AA; 3044 MW; 73C37190403FA349 CRC64;

Query Match 91.3%; Score 21; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02; Mismatches 3; Conservative 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 13 PHG 15

RESULT 9
SODC_OLEEU STANDARD; PRT; 29 AA.
AC P80740;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Allergen Ole e 5) (Ole e V) (Fragment).
OS Olea europaea (Common olive).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Oleaceae; Olea.
OX NCBI_TaxID=4146;
RN [1]
RP SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=98160390; PubMed=9500754;
RA Boluda L., Alonso C., Fernandez-Caldas E.;
RT "Purification, characterization, and partial sequencing of two new allergens of Olea europaea.";
RL J. Allergy Clin. Immunol. 101:210-216(1998).
CC -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; PARTIAL.
DR PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW Allergen; Antioxidant; Copper; Direct protein sequencing;
KW Metal-binding; Oxidoreductase; Zinc.
FT NON TER 29
SQ SEQUENCE 29 AA; 2973 MW; 836C7A193EDAD71E CRC64;

Query Match 91.3%; Score 21; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02; Mismatches 3; Conservative 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 13 PHG 15

RESULT 10
Q6X6Z8 PRELIMINARY; PRT; 29 AA.
AC Q6X6Z8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Topoisomerase (Fragment).
GN Name=parC;
OS Enterococcus gallinarum.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1353;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35038;
RX PubMed=14769469; DOI=10.1016/S0378-1097(03)00929-7;
RA Petersen A., Jensen L.B.;
RT "Analysis of gyrA and parC mutations in enterococci from environmental
RT samples with reduced susceptibility to ciprofloxacin.";
RL FEMS Microbiol. Lett. 231:73-76(2004).
DR EMBL; AY252100; AAP84069.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisoIV.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR ProDom; PD000742; DNA_topoisoIV; 1.
KW Isomerase.
FT NON_TER 1 1
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3180 MW; 88F3936A00740CCD CRC64;

Query Match 91.3%; Score 21; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 13 PHG 15

RESULT 11
Q6X701
ID Q6X701 PRELIMINARY; PRT; 29 AA.
AC Q6X701;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Topoisomerase (Fragment).
GN Name=parC;
OS Enterococcus flavescens.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=37735;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CECT 4481;
RX PubMed=14769469; DOI=10.1016/S0378-1097(03)00929-7;
RA Petersen A., Jensen L.B.;
RT "Analysis of gyrA and parC mutations in enterococci from environmental
RT samples with reduced susceptibility to ciprofloxacin.";
RL FEMS Microbiol. Lett. 231:73-76(2004).
DR EMBL; AY252097; AAP84066.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisoIV.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR ProDom; PD000742; DNA_topoisoIV; 1.
KW Isomerase.
FT NON_TER 1 1
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3180 MW; 88F3936A00740CCD CRC64;

Query Match 91.3%; Score 21; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 13 PHG 15

RESULT 12
Q8R398
ID Q8R398 PRELIMINARY; PRT; 29 AA.
AC Q8R398;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Ptpns1 protein (Fragment).
GN Name=Ptpns1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC025886; AAH25886.1; -.
DR MGD; MGI:108563; Ptpns1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR GO; GO:0045309; F:phosphoprotein amino acid binding; IPI.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007015; P:actin filament organization; IMP.
DR GO; GO:0006928; P:cell motility; IMP.
DR GO; GO:0007160; P:cell-matrix adhesion; IMP.
DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; IMP.
DR GO; GO:0006911; P:phagocytosis, engulfment; IDA.
DR GO; GO:0006910; P:phagocytosis, recognition; IDA.
DR GO; GO:0050766; P:positive regulation of phagocytosis; IDA.
FT NON_TER 1 1
SQ SEQUENCE 29 AA; 3010 MW; F17E24D2CC31E6D1 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 18 PHG 20

RESULT 17

Q7M553 ID Q7M553 PRELIMINARY; PRT; 32 AA.
AC Q7M553;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribosomal protein S4.eR (Fragment).
OS Halobacterium salinarium (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE.
RA Yaguchi M., Visentin L.P., Zuker M., Matheson A.T., Roy C.,
RA Strom A.R.;
RT "Amino-terminal sequences of ribosomal proteins from the 30S subunit
of archaeobacterium Halobacterium cutirubrum.";
RL Submitted (DEC-1990) to the PIR data bank.
DR PIR; S11610; S11610.
FT NON_TER 1 1
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3507 MW; DC612E3DC957273A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 30 PHG 32

RESULT 18

Q6QLS6 ID Q6QLS6 PRELIMINARY; PRT; 32 AA.
AC Q6QLS6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rhodococcus fascians.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-B-15096;
RA Dorado G., Roldan J.M., Leshner J.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AY534865; AAS45278.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3463 MW; BCA73D0B5B1D338A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 26 PHG 28

RESULT 19

Q736U2 ID Q736U2 PRELIMINARY; PRT; 32 AA.
AC Q736U2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BCE2808;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
adaptations and a large plasmid related to Bacillus anthracis pXO1.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017273; AAS41720.1; -.
DR TIGR; BCE2808; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 32 AA; 3635 MW; 6866784DFC0EADA3 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 14 PHG 16

RESULT 20

O02832 ID O02832 PRELIMINARY; PRT; 33 AA.
AC O02832;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Huntingtin protein (Fragment).
GN Name=IT15;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326790; PubMed=8766138;
RA Pecheux C., Gall A.L., Kaplan J.C., Dode C.;
RT "Sequence analysis of the CAG triplet repeats region in the Huntington
disease gene (IT15) in several mammalian species.";
RL Ann. Genet. 39:81-86(1996).
DR EMBL; S83377; AAB50771.1; -.
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3393 MW; 44F558CE5636722A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 7 PHG 9

RESULT 21

O66384 ID O66384 PRELIMINARY; PRT; 33 AA.
AC O66384;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)

Db ||||
6 PHG 8

RESULT 25
HEMO-CHICK
ID HEMO-CHICK STANDARD; PRT; 34 AA.
AC P20057;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hemopexin (Fragment).
GN Name=HPX;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=88339942; PubMed=3421961;
RA Wellner D.; Cheng K.C.; Mueller-Eberhard U.;
RT "N-terminal amino acid sequences of the hemopexins from chicken, rat
and rabbit.";
RL Biochem. Biophys. Res. Commun. 155:622-625(1988).
CC -!- FUNCTION: Binds heme and transports it to the liver for breakdown
and iron recovery, after which the free hemopexin returns to the
circulation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
DR PIR; C31514; C31514.
DR InterPro; IPR000585; Hemopexin.
DR PROSITE; PS00024; HEMOPEXIN; PARTIAL.
KW Direct protein sequencing; Glycoprotein; Heme; Plasma; Repeat;
KW Transport.
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 3598 MW; 07A2EEAC27386C57 CRC64;

Query Match 91.3%; Score 21; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 16 PHG 18

RESULT 26
Q73RE8
ID Q73RE8 PRELIMINARY; PRT; 34 AA.
AC Q73RE8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=TDE0141;
OS Treponema denticola.
OC Bacteria; Spirochaetes.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gebregeorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
RA Shatsman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstock G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
with other spirochete genomes.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017246; AAS10639.1; -.
DR TIGR; TDE0141; -.
KW Complete proteome.
SQ SEQUENCE 34 AA; 3643 MW; EC84B7F1133531C4 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 28 PHG 30

RESULT 27
Q6LBR0
ID Q6LBR0 PRELIMINARY; PRT; 35 AA.
AC Q6LBR0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE T10 class I MHC gene (exon 5) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Domesticus;
RX MEDLINE=90063453; PubMed=2584927;
RA Brorson K.A., Hunt S.W., Hunkapiller T., Sun H.Y., Cheroutre H.,
RA Nickerson D.A., Hood L.;
RT "Comparison of exon 5 sequences from 35 class I genes of the BALB/c
mouse.";
RT J. Exp. Med. 170:1837-1858(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Domesticus;
RA Brorson K.;
RL Submitted (AUG-1989) to the EMBL/GenBank/DBJ databases.
DR EMBL; X16216; CAE82028.1; -.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3770 MW; 695C1D11D6C7A052 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|||
Db 8 PHG 10

RESULT 28
Q8CLW4
ID Q8CLW4 PRELIMINARY; PRT; 35 AA.
AC Q8CLW4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical.
GN OrderedLocusNames=y0064;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AE013607; AAM83659.1; -.
KW Hypothetical protein.
SQ SEQUENCE 35 AA; 3936 MW; 10744A0FDF6AC04B CRC64;

Query Match 91.3%; Score 21; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 33 PHG 35
|||

RESULT 29
Q8ZWQ4 PRELIMINARY; PRT; 36 AA.
ID Q8ZWQ4
AC Q8ZWQ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PaREP2a.
GN OrderedLocusNames=PAE1672;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009833; AAL63646.1; -.
KW Complete proteome.
SQ SEQUENCE 36 AA; 4060 MW; 43CC1A624F2822C5 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 10 PHG 12
|||

RESULT 30
Q9URE9 PRELIMINARY; PRT; 36 AA.
ID Q9URE9
AC Q9URE9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Multidrug resistance P-glycoprotein (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93368189; PubMed=8361216;
RA Allikmets R., Gerrard B., Stewart C., White M., Dean M.;
RT "Identification of P-glycoprotein/multidrug resistance genes from
model organisms.";

RL Leukemia 7:13-13(1993).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC_tran; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
KW ATP-binding.
SQ SEQUENCE 36 AA; 3753 MW; CDB46C744DC8610D CRC64;

Query Match 91.3%; Score 21; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 2 PHG 4
|||

RESULT 31
Q7TNQ0 PRELIMINARY; PRT; 37 AA.
ID Q7TNQ0
AC Q7TNQ0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC056177; AAH56177.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 3977 MW; 5D7166A450809A2B CRC64;

Query Match 91.3%; Score 21; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
Db 3 PHG 5
|||

Db 3 PHG 5

RESULT 32

Q6X700 PRELIMINARY; PRT; 38 AA.

ID Q6X700

AC Q6X700

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Topoisomerase (Fragment).

GN Name=parC;

OS Enterococcus casseliflavus.

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

OX NCBI_TaxID=37734;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 25788;

RX PubMed=14769469; DOI=10.1016/S0378-1097(03)00929-7;

RA Petersen A., Jensen L.B.;

RT "Analysis of gyrA and parC mutations in enterococci from environmental samples with reduced susceptibility to ciprofloxacin.";

RL FEMS Microbiol. Lett. 231:73-76(2004).

DR EMBL; AY252098; AAP84067.1; -.

DR HSSP; P09097; IAB4.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.

DR GO; GO:0016853; F:isomerase activity; IEA.

DR GO; GO:0006265; P:DNA topological change; IEA.

DR InterPro; IPR002205; DNA_topoisoIV.

DR Pfam; PF00521; DNA_topoisoIV; 1.

DR ProDom; PD000742; DNA_topoisoIV; 1.

KW Isomerase.

FT NON_TER 1 1

FT NON_TER 38 38

SQ SEQUENCE 38 AA; 4357 MW; CEB751F1C22440E8 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 13 PHG 15

RESULT 33

Q8VJV6 PRELIMINARY; PRT; 38 AA.

ID Q8VJV6

AC Q8VJV6;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=MT1909;

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RX MEDLINE=22206494; PubMed=12218036;

RX DOI=10.1128/JB.184.19.5479-5490.2002;

RA Fleischohn R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H., Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;

RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";

RL J. Bacteriol. 184:5479-5490(2002).

DR EMBL; AE000516; AAK46180.1; -.

DR TIGR; MT1909; -.

KW Hypothetical protein.

SQ SEQUENCE 38 AA; 4273 MW; 461CADD7180DB359 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 10 PHG 12

RESULT 34

Q81TC9 PRELIMINARY; PRT; 38 AA.

ID Q81TC9

AC Q81TC9; Q6KVF5;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=BA1355, GBAA1355;

OS Bacillus anthracis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1392;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Ames / isolate Porton;

RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;

RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T., Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R., Holtzapple E.K., Okstad O.A., Helgason E., Rikstone J., Wu M., Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L., DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H., Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D., Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F., Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C., Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O., Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolstoe A.-B., Fraser C.M.;

RT "The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria.";

RL Nature 423:81-86(2003).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Ames / isolate 0581;

RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B., Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L., Fraser C.M.;

RT "Bacillus anthracis comparative genomics.";

RL Submitted (MAY-2004) to the EMBL/GenBank/DBDJ databases.

DR EMBL; AE017028; AAP25301.1; -.

DR EMBL; AE017334; AAT30451.1; -.

DR TIGR; BA1355; -.

DR TIGR; GBAA1355; -.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 38 AA; 4456 MW; 965D6E445A9BE25C CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 21 PHG 23

RESULT 35

Q9RUH4 PRELIMINARY; PRT; 38 AA.

ID Q9RUH4

AC Q9RUH4;

```
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein DR1412.
GN OrderedLocNames=DR1412;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001986; AAF10978.1; -.
DR PIR; G75398; G75398.
DR TIGR; DR1412; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 38 AA; 3891 MW; 0EE8EFA6AFA325A4 CRC64;

Query Match          91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
31 PHG 33

RESULT 36
Q6EES4
ID Q6EES4 PRELIMINARY; PRT; 39 AA.
AC Q6EES4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE GRINL1A combined protein Gcom12 (Fragment).
GN Name=GRINL1A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX PubMed=15233991; DOI=10.1016/j.ygeno.2004.04.004;
RA Roginski R.S., Mohan Raj B.K., Birditt B., Rowen L.;
RT "The human GRINL1A gene defines a complex transcription unit, an
RT unusual form of gene organization in eukaryotes.";
RL Genomics 84:265-276(2004).
DR EMBL; AY350748; AAQ82541.1; -.
DR NON_TER 1
SQ SEQUENCE 39 AA; 4393 MW; 530FE3E4CC04D4A5 CRC64;

Query Match          91.3%; Score 21; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
22 PHG 24

RESULT 37
```

```
Q9NC67
ID Q9NC67 PRELIMINARY; PRT; 39 AA.
AC Q9NC67;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sphingosine-1-phosphate lyase (Fragment).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20431894; PubMed=10974109;
RA Li G., Alexander H., Schneider N., Alexander S.;
RT "Molecular basis for resistance to the anticancer drug cisplatin in
RT Dictyostelium.";
RL Microbiology 146:2219-2227(2000).
DR EMBL; AF233610; AAF97870.1; -.
DR DictyBase; DDB0214888; sglA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR010916; TONB Box N.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Lyase.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4140 MW; BBS528F43199727A CRC64;

Query Match          91.3%; Score 21; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
32 PHG 34

RESULT 38
Q6X702
ID Q6X702 PRELIMINARY; PRT; 39 AA.
AC Q6X702;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gyrase (Fragment).
GN Name=gyrA;
OS Enterococcus flavescens.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=37735;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CECT 4481;
RX PubMed=14769469; DOI=10.1016/S0378-1097(03)00929-7;
RA Petersen A., Jensen L.B.;
RT "Analysis of gyrA and parC mutations in enterococci from environmental
RT samples with reduced susceptibility to ciprofloxacin.";
RL FEMS Microbiol. Lett. 231:73-76(2004).
DR EMBL; AY252096; AAP84065.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisoIV.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR ProDom; PD000742; DNA_topoisoIV; 1.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4312 MW; CDB3CDD0246265F3 CRC64;

Query Match          91.3%; Score 21; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 3 PHG 5
 |||
Db 31 PHG 33
 0; Gaps 0;

Query Match 91.3%; Score 21; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
 |||
Db 13 PHG 15

Search completed: June 15, 2005, 14:21:22
Job time : 85.5 secs

RESULT 39
Q8ZTN8 PRELIMINARY; PRT; 40 AA.
AC Q8ZTN8;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE PaREP2a.
GN OrderedLocusNames=PAE3166;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009915; AAL64721.1; -.
KW Complete proteome.
SQ SEQUENCE 40 AA; 4604 MW; B6C670022B6F64E7 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
 |||
Db 14 PHG 16

RESULT 40
P95768 PRELIMINARY; PRT; 40 AA.
AC P95768;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE GryA (Fragment).
GN Name=gyrA;
OS Staphylococcus haemolyticus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1283;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96254595; PubMed=8849234;
RA Yonezawa M., Takahata M., Banzawa-Futakuchi N., Matsubara N.,
RA Watanabe Y., Narita H., Matsunaga T., Igarashi H., Kawahara M.,
RA Onodera S.;
RT "DNA gyrase gyrA mutations in quinolone-resistant clinical isolates of
RT Staphylococcus haemolyticus.";
RL Antimicrob. Agents Chemother. 40:1065-1066(1996).
DR EMBL; D78568; BAA11412.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisomerase.
DR Pfam; PF00521; DNA_topoisomerase.
DR ProDom; PD000742; DNA_topoisomerase.
FT NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4533 MW; A4C6842027AA7812 CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-8

Perfect score: 37

Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	37	100.0	5	5	ABB79809	Abb79809 Histidine
2	37	100.0	5	8	ADH10415	Adh10415 Rabbit HP
3	37	100.0	13	5	AAU10991	Aau10991 Cathepsin
4	37	100.0	33	4	AAE10907	Aae10907 Human bra
5	37	100.0	42	4	AAO07477	Aao07477 Human pol
6	37	100.0	97	4	AAG73663	Aag73663 Human col
7	37	100.0	101	8	ADH10412	Adh10412 Rabbit HP
8	37	100.0	137	8	ADM16706	Adm16706 PRM-P5-F1
9	37	100.0	148	5	ABB79807	Abb79807 Rabbit hi
10	37	100.0	148	5	ABB79806	Abb79806 Human his
11	37	100.0	148	8	ADH10411	Adh10411 Human HPR
12	37	100.0	151	4	AAM21476	Aam21476 Peptide #
13	37	100.0	151	4	ABB43818	Abb43818 Peptide #
14	37	100.0	151	4	AAM37725	Aam37725 Peptide #
15	37	100.0	151	4	ABB26753	Abb26753 Protein #
16	37	100.0	151	4	AAM77540	Aam77540 Human bon
17	37	100.0	151	4	AAM64782	Aam64782 Human bra
18	37	100.0	151	4	ABG59189	Abg59189 Human liv
19	37	100.0	151	5	ABG46564	Abg46564 Human pep
20	37	100.0	198	4	ABB71093	Abb71093 Drosophil
21	37	100.0	248	4	AAU55932	Aau55932 Propionib
22	37	100.0	248	6	ABM52451	Abm52451 Propionib
23	37	100.0	259	6	ABU35012	Abu35012 Protein e
24	37	100.0	264	8	ADL05482	Adl05482 M. catarr
25	37	100.0	273	2	AAY32079	Aay32079 Human gap

26	37	100.0	273	5	ADI17057	Adi17057 Human NOV
27	37	100.0	273	5	ADI17060	Adi17060 Human NOV
28	37	100.0	273	6	ABR48232	Abr48232 Human bla
29	37	100.0	273	6	ABU56449	Abu56449 Lung canc
30	37	100.0	273	6	ABU56673	Abu56673 Lung canc
31	37	100.0	273	7	ADN39026	Adn39026 Cancer/an
32	37	100.0	273	7	ADN39932	Adn39932 Cancer/an
33	37	100.0	323	4	ABB69468	Abb69468 Drosophil
34	37	100.0	326	8	ADN46756	Adn46756 Thermococ
35	37	100.0	327	4	ABB66932	Abb66932 Drosophil
36	37	100.0	350	8	ADS30838	Ads30838 Bacterial
37	37	100.0	402	4	ABB71820	Abb71820 Drosophil
38	37	100.0	419	4	ABB59479	Abb59479 Drosophil
39	37	100.0	466	7	ABO77799	Abo77799 Pseudomon
40	37	100.0	525	5	ABB79804	Abb79804 Human his
41	37	100.0	525	8	ADE76897	Ade76897 Human pro
42	37	100.0	525	8	ADH10409	Adh10409 Human HPR
43	37	100.0	525	8	ADL12335	Adl12335 Human ste
44	37	100.0	526	5	ABB79805	Abb79805 Rabbit hi
45	37	100.0	526	8	ADH10410	Adh10410 Rabbit HP

ALIGNMENTS

RESULT 1
ABB79809
ID ABB79809 standard; peptide; 5 AA.

AC ABB79809;
XX
DT 25-NOV-2002 (first entry)
XX
DE Histidine proline rich glycoprotein pentapeptide.
XX
KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW neurotropic; neuroprotective; antiparkinsonian.
XX
OS Synthetic.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
DR P-ESDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 2; Page 67; 82pp; English.
XX

CC The present sequence is a specific example of claimed anti-angiogenic
CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
CC peptides comprise: the histidine-proline-rich (H/P) domain of human
CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808, such as the present peptide, or its variant
CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
CC N- or C-terminus. Also claimed are: chemically synthesised or

CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX Sequence 5 AA;

Query Match 100.0%; Score 37; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

|||||

Db 1 HHPHG 5

RESULT 2

ADH10415

ID ADH10415 standard; peptide; 5 AA.

XX ADH10415;

XX 11-MAR-2004 (first entry)

XX Rabbit HPRG protein H/P rich domain repeat fragment.

DE Tpm; tropomyosin; antiangiogenic receptor;

XX histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;

KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;

KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;

KW angiogenesis; gene therapy; rabbit.

XX Oryctolagus cuniculus.

OS WO2003077872-A2.

XX 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

PF 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

PA

XX Mc Crae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.

XX Example 5; SEQ ID NO 29; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

Query Match 100.0%; Score 37; DB 8; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

|||||

Db 1 HHPHG 5

RESULT 3

AAU10991

ID AAU10991 standard; peptide; 13 AA.

XX AAU10991;

XX 26-MAR-2002 (first entry)

XX Cathepsin D substrate.

XX Contrast agent; cancer; cardiovascular disease; inflammation; stroke;

KW central nervous system disorder; infection; cardiac failure; embolism;

KW myocardial infarction; atherosclerosis; thrombosis; aneurysm;

KW haemorrhage; Alzheimer's disease; multiple sclerosis; bone disease;

KW osteoporosis; apoptosis; necrosis; cathepsin D.

XX Synthetic.

XX Key Modified-site 1 Location/Qualifiers

FT /note= "Pn216-succinyl. Pn216 = Bis[(1,1-dimethyl-2-N-

FT hydroxyimine propyl) aminoethyl]-2-aminoethyl amine"

FT Modified-site 9

FT /note= "Cys(Et)"

XX WO200189584-A2.

PN 29-NOV-2001.

XX 23-MAY-2001; 2001WO-NO000215.

PF 23-MAY-2000; 2000NO-00002644.

XX (NYCO-) NYCOMED IMAGING AS.

PA Klaveness J, Tolleshaug H;

XX WPI; 2002-097616/13.

XX Contrast agents susceptible of changing pharmacodynamic and/or pharmacokinetic properties upon enzymatic activity, useful in diagnosis of cancer, cardiovascular diseases and inflammation in humans and animals.

PT Example 2; Page 53; 77pp; English.

XX The invention relates to a contrast agent substrate (I) susceptible of changing pharmacodynamic and/or pharmacokinetic properties upon the influence of enzymatic activity. The contrast agent substrate is useful for detection of tissue or cells with abnormal metabolic activity, for identification and/or diagnosis of cancer, cardiovascular diseases, diseases on the central nervous system, inflammations, or infections and detection of an area of a disease of abnormal enzymatic activity, where the substrate is administered to human or animal body and a contrast agent signal is detected as a result of the contrast agent changing pharmacodynamic and/or pharmacokinetic properties. The substrate is useful for manufacturing a medicament for detecting an area of disease of abnormal enzymatic activity. The contrast agent is useful in diagnosis of cardiac failure, myocardial infarction, atherosclerosis, thrombosis, embolism, aneurysms, stroke, haemorrhage, central nervous system diseases, preferably Alzheimer's disease or multiple sclerosis, bone diseases such as osteoporosis, viral infections, and for identification of apoptosis and necrosis. The metabolically sensitive contrast agents are more sensitive to pathology than morphological contrast agents. As abnormal enzymatic activity is an early sign of disease/condition, the contrast agents have a potential for diagnosing disease at an early stage, which in many clinical situations are important for the outcome of the treatment. The agents are very sensitive to treatment and can be used to follow up treatment. The present sequence represents the cathepsin D substrate, used in the method of the invention

XX Sequence 13 AA;

SQ Query Match 100.0%; Score 37; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||

Db 2 HHPHG 6

RESULT 4
AAE10907

ID AAE10907 standard; peptide; 33 AA.

XX AAE10907;

AC 18-DEC-2001 (first entry)

XX Human brain sodium channel 4 protein (BNAc4) C-terminal peptide.

DE Human; brain sodium channel 4 protein; BNAc4; cerebroprotective; acid-sensing ion channel; ASIC4; neuromodulation; vasotropic; neurotransmission; ischaemia; neurodegeneration; Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis; antiinflammatory;

KW pain; Parkinsonism; inflammatory condition; cerebellar ataxia; complex regional pain syndrome; osteoarthritis; rheumatoid arthritis; ulcer; neurodegenerative disease; asthma; irritable bowel syndrome; chronic obstructive pulmonary disease; central nervous system disease; psoriasis; intestinal tract disease; gene therapy; neuroprotective; nootropic; analgesic; osteopathic; cancer.

XX Homo sapiens.

OS

XX Key Location/Qualifiers

FH Misc-difference 2 /label= Lys, Arg

FT Misc-difference 5 /label= Ala, Val

XX WO200166125-A1.

XX 13-SEP-2001.

PD 13-FEB-2001; 2001WO-US0004525.

XX 03-MAR-2000; 2000US-00518959.

PR (ORTH) ORTHO-MCNEIL PHARM INC.

XX Dubin AE, Erlander MG, Huvar R, Pyati J;

XX WPI; 2001-589913/66.

XX Novel human brain sodium channel 4 protein, a member of acid-sensing ion channel family, for identifying modulators comprises measuring the effect of the modulator on the protein.

PS Example 7; Page 60; 94pp; English.

XX The invention relates to human brain sodium channel 4 protein (BNAc4), a member of acid-sensing ion channel family (ASIC4). Human BNAc4 is useful for identifying compounds that modulate its activity, by measuring the effect of the modulator on the protein, BNAc-dependent cytotoxicity or BNAc-dependent activation of ion flux. Compounds discovered by screening test compounds using human BNAc4 protein is useful for treating a patient in need of a treatment for a condition mediated by human BNAc4. BNAc4 antibody is useful for blocking the activity of the polypeptide. BNAc4 plays an important role in neuromodulation, neurotransmission, pain, ischaemia and neurodegeneration underlying diseases including Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis, cerebellar ataxias and Parkinsonism. Modulators identified using this protein are useful as therapeutic agents, for the treatment of ischaemia inflammatory conditions and for use as analgesics for intractable pain, complex regional pain syndrome, arthritis (osteoarthritis, rheumatoid), ulcers, neurodegenerative diseases, asthma, chronic obstructive pulmonary disease, irritable bowel syndrome, psoriasis, central nervous system diseases, diseases of the intestinal tract, abnormal proliferation and cancer especially in the digestive system and female gonads, ulcer, liver disease, control of viscera innervated by the dorsal root ganglia and to diagnose or treat any disorder related to abnormal expression of the human BNAc4 polypeptides. BNAc4 DNA and its portions are useful for isolating homologues of the DNA molecules, identifying and isolating genomic equivalents of the DNA molecules and identifying, detecting or isolating mutant forms of the DNA molecules. BNAc4 DNA is used in gene therapy. The present sequence is human BNAc4 C-terminal peptide. There were 2 bases in the BNAc4 coding sequence that were consistently different in one of 10 independent cDNA library clones. These keto and pyrimidine differences caused an L to become R and A to V in the C-terminal peptide

XX Sequence 33 AA;

SQ Query Match 100.0%; Score 37; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

Db 18 HHPHG 22

RESULT 5
AAO07477
ID AAO07477 standard; protein; 42 AA.
XX
AC AAO07477;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 21369.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI87408.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 21369; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 42 AA;

Query Match 100.0%; Score 37; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. NO. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 38 HHPHG 42

RESULT 6
AAG73663
ID AAG73663 standard; protein; 97 AA.
XX
AC AAG73663;
XX

DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:4427.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH33094.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 11; Page 6244-6245; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 97 AA;

Query Match 100.0%; Score 37; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 53 HHPHG 57

RESULT 7
ADH10412
ID ADH10412 standard; protein; 101 AA.
XX
AC ADH10412;
XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein H/P rich domain.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;

KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Oryctolagus cuniculus.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
DR
XX
PI New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Claim 6; SEQ ID NO 24; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 37; DB 8; Length 101;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 13 HHPHG 17

RESULT 8
ADM16706
ID ADM16706 standard; protein; 137 AA.
XX
AC ADM16706;
XX
DT 17-JUN-2004 (first entry)
XX
DE PRM-P5-F12 protein.

XX Phlebotomus ariasi; Phlebotomus perniciosus; Vaccine;
KW Leishmania infection.
XX
OS Phlebotomus arisasi.
XX
PN WO2004027041-A2.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029833.
XX
PR 19-SEP-2002; 2002US-0412327P.
PR 12-NOV-2002; 2002US-0425852P.
XX
PA (MERI-) MERIAL LTD.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Valenzuela JG, Ribiero JMC, Kanhawi S, Belkaid Y, Fischer L;
PI Audonnet J, Milward F;
XX
DR WPI; 2004-295410/27.
DR N-PSDB; ADM16707.
XX
PT New substantially purified salivary Phlebotomus ariasi or Phlebotomus
PT perniciosus polypeptide, useful for inhibiting, treating or preventing
PT Leishmania infection in a subject.
XX
PS Claim 1; SEQ ID NO 29; 200pp; English.
XX
CC The present invention relates to a substantially purified salivary
CC Phlebotomus ariasi or Phlebotomus perniciosus polypeptide. The
CC composition comprising the polypeptide or the nucleic acid encoding the
CC polypeptide is useful for manufacture of a medicament. The polypeptides
CC and nucleic acids are useful for inducing an immune response to a P.
CC ariasi or P. perniciosus polypeptide. It is also useful for inhibiting a
CC symptom of a Leishmania infection or preventing a Leishmania infection in
CC a subject. The present sequence represents a purified Phlebotomus ariasi
CC protein.
XX
SQ Sequence 137 AA;

Query Match 100.0%; Score 37; DB 8; Length 137;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 61 HHPHG 65

RESULT 9
ABB79807
ID ABB79807 standard; protein; 148 AA.
XX
AC ABB79807;
XX
DT 25-NOV-2002 (first entry)
XX
DE Rabbit histidine proline rich glycoprotein His/Pro-rich domain.
XX
KW Histidine proline rich glycoprotein; HPGP; rabbit; antiangiogenic;
KW cytosstatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
XX
OS Oryctolagus cuniculus.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.

XX PR 14-FEB-2001; 2001US-0268370P.
XX PA (ATTE-) ATTENUON LLC.
XX PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX XX
DR WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 13; 82pp; English.
XX
CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of rabbit anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79805). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain (see ABB79806) of human
CC HPRG; an H/P domain of rabbit HPRG; a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: Chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
XX immunassays

SQ Sequence 148 AA;
Query Match 100.0%; Score 37; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHPHG 5
Dbb 35 HHPHG 39
RESULT 10
ABB79806
ID ABB79806 standard; protein; 148 AA.

XX ABB79806;
AC
XX
DT 25-NOV-2002 (first entry)
XX
DE Human histidine proline rich glycoprotein His/Pro-rich domain.
XX
KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian.
XX
OS Homo sapiens.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX (ATTE-) ATTENUON LLC.
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI; 2002-666989/71.
DR P-PSDB; ABN84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 13; 82pp; English.
XX
CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of human anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79804). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain of human HPRG; an H/P
CC domain of rabbit HPRG (see ABB12345); a variant of these that is capable
CC of inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: Chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or peptide
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
XX immunassays

CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ Sequence 148 AA;
Query Match 100.0%; Score 37; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 35 HHPHG 39
RESULT 11
ADH10411
ID ADH10411 standard; protein; 148 AA.
XX
AC ADH10411;
XX
DT 11-MAR-2004 (first entry)
DE Human HPRG protein H/P rich domain.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-US008060.
XX
PR 15-MAR-2002; 2002US-0364047P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mccrae K, Donate F, Juarez J, Mazar AP;
XX
XX WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Claim 6; SEQ ID NO 23; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing

CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a human HPRG-
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 148 AA;
Query Match 100.0%; Score 37; DB 8; Length 148;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 35 HHPHG 39
RESULT 12
AAM21476
ID AAM21476 standard; protein; 151 AA.
XX
AC AAM21476;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #7910 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488901/53.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 27; SEQ ID NO 26302; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs; see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 128 HHPHG 132

RESULT 13
ABB43818
ID ABB43818 standard; peptide; 151 AA.
XX
AC ABB43818;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #11324 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 36453; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 128 HHPHG 132

RESULT 14
AAM37725
ID AAM37725 standard; protein; 151 AA.
XX
AC AAM37725;

XX 17-OCT-2001 (first entry)
DT
XX Peptide #11762 encoded by probe for measuring placental gene expression.
DE
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
KW
XX Homo sapiens.
OS
XX WO200157272-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000663.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
PT
XX Claim 27; SEQ ID NO 37994; 654pp; English.
PS
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 128 HHPHG 132

RESULT 15
ABB26753
ID ABB26753 standard; protein; 151 AA.
XX
AC ABB26753;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #8752 encoded by probe for measuring heart cell gene expression.
DE
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
OS
XX WO200157274-A2.
PN
XX 09-AUG-2001.
PD
XX

```
PF 30-JAN-2001; 2001WO-US0000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 28523; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db |||||
128 HHPHG 132

RESULT 16
AAM77540
ID AAM77540 standard; protein; 151 AA.
XX
AC AAM77540;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 37846.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-488899/53.
Single exon nucleic acid probes for analyzing gene expression in human
hearts.
Claim 15; SEQ ID NO 28523; 530pp; English.
The present invention relates to single exon nucleic acid probes for
measuring human gene expression in a sample derived from human heart (see
ABA21535-ABA41305). The present sequence is a protein encoded by one such
probe. The probes may be used for predicting, measuring and displaying
gene expression in samples derived from the human heart via microarrays.
By measuring gene expression, the probes are useful for predicting,
diagnosing, grading, staging, monitoring and prognosing diseases of the
human heart and vascular system e.g. cardiovascular disease,
hypertension, cardiac arrhythmias and congenital heart disease. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences
Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db |||||
128 HHPHG 132

RESULT 17
AAM64782
ID AAM64782 standard; protein; 151 AA.
XX
AC AAM64782;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 36887.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-483446/52.
Single exon nucleic acid probes for analyzing gene expression in human
brains.
Example 4; SEQ ID NO 36887; 650pp + Sequence Listing; English.
The present invention provides a number of single exon nucleic acid
probes which are derived from genomic sequences expressed in the human
brain. They can be used to measure gene expression in brain cell samples,
```

CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention

XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 128 HHPHG 132

RESULT 18
ABG59189
ID ABG59189 standard; peptide; 151 AA.

XX
AC ABG59189;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 37837.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.

XX
OS Homo sapiens.

XX
PN WO200157273-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000664.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

XX
PR 27-SEP-2000; 2000US-0236359P.

XX
PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-48898/53.

XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.

XX
PS Claim 27; SEQ ID NO 37837; 658pp; English.

XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 128 HHPHG 132

RESULT 19
ABG46564
ID ABG46564 standard; peptide; 151 AA.

XX
AC ABG46564;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 36229.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.

XX
OS Homo sapiens.

XX
PN WO200186003-A2.

XX
PD 15-NOV-2001.

XX
PF 30-JAN-2001; 2001WO-US000665.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

XX
PR 21-SEP-2000; 2000US-0234687P.

XX
PR 27-SEP-2000; 2000US-0236359P.

XX
PR 04-OCT-2000; 2000GB-00024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2002-114183/15.

XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.

XX
PS Claim 27; SEQ ID NO 36229; 634pp; English.

XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,

comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karageners syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Qy 1 HHPHG 5
Db 128 HHPHG 132

Query Match 100.0%; Score 37; DB 5; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Qy 1 HHPHG 5
Db 153 HHPHG 157

Query Match 100.0%; Score 37; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 21
AAU55932

ID AAU55932 standard; protein; 248 AA.

AC AAU55932;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #16828.

SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

DR N-PSDB; AAS59573.

PT Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX Example 1; SEQ ID NO 17127; 1069pp; English.

Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

Disclosure; SEQ ID NO 40071; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and

CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 248 AA;

Query Match 100.0%; Score 37; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 142 HHPHG 146

RESULT 22
ABM52451
ID ABM52451 standard; protein; 248 AA.
XX
AC ABM52451;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #17127.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64502.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 17127; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 248 AA;

Query Match 100.0%; Score 37; DB 6; Length 248;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 142 HHPHG 146

RESULT 23
ABU35012
ID ABU35012 standard; protein; 259 AA.
XX
AC ABU35012;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #20539.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Moraxella catarrhalis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA38882.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 62936; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 259 AA;
Query Match 100.0%; Score 37; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 229 HHPHG 233

RESULT 24
ADL05482
ID ADL05482 standard; protein; 264 AA.

XX
AC ADL05482;
XX
DT 06-MAY-2004 (first entry)
XX
DE M. catarrhalis protein #1248.
XX
KW Moraxella catarrhalis; infection.
XX
OS Moraxella catarrhalis.
XX
PN US6673910-B1.
XX
PD 06-JAN-2004.
XX
PF 04-APR-2000; 2000US-00540236.
XX
PR 08-APR-1999; 99US-0128416P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
WPI; 2004-178127/17.
DR N-PSDB; ADL03562.

XX
PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.

XX Disclosure; SEQ ID NO 3168; 429pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents the amino acid
CC sequence of a M. catarrhalis protein.

XX
SQ Sequence 264 AA;
Query Match 100.0%; Score 37; DB 8; Length 264;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 230 HHPHG 234

RESULT 25
AAY32079
ID AAY32079 standard; protein; 273 AA.

XX
AC AAY32079;
XX
DT 17-JAN-2000 (first entry)
XX
DE Human gap junction protein beta-4.
XX
KW Gap junction protein beta-4; human; neurological disease;
KW epidermal disease; deafness; cataract; AIDS; therapy; diagnosis; vaccine.
XX
OS Homo sapiens.
XX
PN WO9951738-A1.
XX
PD 14-OCT-1999.
XX
PF 03-APR-1998; 98WO-CN0000055.
XX
PR 03-APR-1998; 98WO-CN0000055.
XX
PA (UYHU-) UNIV HUNAN MEDICAL.
XX
PI Xia J, Pan Q, Liu C, Zheng D, Xie W;
XX
DR WPI; 1999-611041/52.
DR N-PSDB; AAZ20259.
XX
PT Novel polypeptides and polynucleotides useful for treating and preventing
PT e.g. neurological or epidermal diseases.

XX Claim 1; Page 25; 33pp; English.
XX
CC The present sequence represents human gap junction protein beta-4, a
CC protein that has homology to connexins. The invention relates to gap
CC junction protein beta-4 polypeptides and polynucleotides (see also
CC AAZ20259), recombinant materials and methods for their production. It
CC also relates to methods for using such polypeptides and polynucleotides
CC for the treatment of neurological diseases, epidermal diseases, deafness,
CC cataract, and AIDS, and for identifying agonists, antagonists and/or
CC inhibitors useful in treating conditions associated with gap junction
CC protein beta-4 imbalance. Also provided are diagnostic assays for
CC detecting diseases associated with inappropriate gap junction protein
CC beta-4 activity or levels, e.g. by analysing for the presence or amount
CC of polypeptide expression in a sample. The polypeptides are also useful
CC for raising antibodies immunospecific to gap junction protein beta-4

XX Sequence 273 AA;

XX
Query Match 100.0%; Score 37; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 229 HHPHG 233

RESULT 26

ADII17057
ID ADII17057 standard; protein; 273 AA.
XX
AC ADII17057;
XX
DT 15-APR-2004 (first entry)
XX
DE Human NOVX protein homologue SeqID 593.
XX
KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
XX
OS Homo sapiens.
XX
PN WO200268649-A2.
XX
PD 06-SEP-2002.
XX
PF 31-JAN-2002; 2002WO-US002785.
XX
PR 31-JAN-2001; 2001US-02653395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 07-FEB-2001; 2001US-0267057P.
PR 08-FEB-2001; 2001US-0267459P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
PI Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2002-706998/76.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Disclosure; SEQ ID NO 593; 1498pp; English.
XX
CC This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cyostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 273 AA;

Query Match 100.0%; Score 37; DB 5; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db |||||
229 HHPHG 233

RESULT 27
ADII17060
ID ADII17060 standard; protein; 273 AA.
XX
AC ADII17060;
XX
DT 15-APR-2004 (first entry)
XX
DE Human NOVX protein homologue SeqID 596.
XX
KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

KW Alzheimer's disease; infection; str.
XX Homo sapiens.
OS WO200268649-A2.
XX PD 06-SEP-2002.
XX 31-JAN-2002; 2002WO-US002785.
PF 31-JAN-2001; 2001US-0265395P.
XX 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 07-FEB-2001; 2001US-0267057P.
PR 08-FEB-2001; 2001US-0267459P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX (CURA-) CURAGEN CORP.
PA Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
XX Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.

XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX Disclosure; SEQ ID NO 596; 1498pp; English.
PS
XX
CC This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 273 AA;
Query Match 100.0%; Score 37; DB 5; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 229 HHPHG 233
RESULT 28
ABR48232
ID ABR48232 standard; protein; 273 AA.
XX
AC ABR48232;
XX 12-JUN-2003 (first entry)
XX Human bladder cancer associated protein sequence SEQ ID NO:185.
DE Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX Homo sapiens.
XX WO2003003906-A2.
XX 16-JAN-2003.
XX 03-JUL-2002; 2002WO-US021338.
PF 03-JUL-2001; 2001US-0302814P.
XX 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA

XX Mack DH, Aziz N;
PI
XX WPI; 2003-201532/19.
DR N-PSDB; ACC51048.
DR
XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
PT
XX
PS Claim 10; Page 293; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 273 AA;

Query Match 100.0%; Score 37; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 229 HHPHG 233

RESULT 29
ABU56449
ID ABU56449 standard; protein; 273 AA.
XX
AC ABU56449;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #42.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
XX WO200286443-A2.
XX
XX 31-OCT-2002.
XX
XX 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
PI

XX WPI; 2003-093161/08.
DR N-PSDB; ABX76170.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 222; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 273 AA;

Query Match 100.0%; Score 37; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 229 HHPHG 233

RESULT 30
ABU56673
ID ABU56673 standard; protein; 273 AA.
XX
AC ABU56673;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #266.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
XX WO200286443-A2.
XX
XX 31-OCT-2002.
XX
XX 18-APR-2002; 2002WO-US012476.
XX
PF 18-APR-2001; 2001US-0284770P.
XX
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
PI

PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
XX N-PSDB; ABX76402.
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 393; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 273 AA;
Query Match 100.0%; Score 37; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db |||||
229 HHPHG 233
RESULT 31
ADN39026
ID ADN39026 standard; protein; 273 AA.
XX
AC ADN39026;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:344.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerability; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
PI WPI; 2003-468649/44.
XX N-PSDB; ADN39025.
DR Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
PS Claim 12; SEQ ID NO 344; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 273 AA;
Query Match 100.0%; Score 37; DB 7; Length 273;
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db |||||
229 HHPHG 233
RESULT 32
ADN39932
ID ADN39932 standard; protein; 273 AA.
XX
AC ADN39932;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C302.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;

KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX vulnerary; gene therapy; vaccine.

OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
PF
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynnne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR N-PSDB; ADN39715.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

XX Claim 12; SEQ ID NO C302; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.

XX Sequence 273 AA;
SQ

Query Match 100.0%; Score 37; DB 7; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
|
|
|
Db 229 HHPHG 233

RESULT 33
ABB69468
ID ABB69468 standard; protein; 323 AA.
XX
AC ABB69468;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 35196.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL13571.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Disclosure; SEQ ID NO 35196; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 323 AA;
SQ

Query Match 100.0%; Score 37; DB 4; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
|
|
|
Db 279 HHPHG 283

RESULT 34
ADN46756
ID ADN46756 standard; protein; 326 AA.
XX
AC ADN46756;
XX
DT 01-JUL-2004 (first entry)
XX
DE Thermococcus kodakaraensis KOD1 protein sequence SeqID634.
XX
KW gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;

KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology.
XX
OS Thermococcus kodakaraensis.
XX
PN WO2004022736-A1.
XX
PD 18-MAR-2004.
XX
PF 29-AUG-2003; 2003WO-IB003597.
XX
PR 30-AUG-2002; 2002JP-00319011.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI Imanaka T, Atomi H;
XX
DR WPI; 2004-257583/24.
XX
PT Method for disrupting targeted gene in genome of organism particularly
PT thermostable bacterium and with genome chips for analysis, applicable in
PT studying gene structure and functions.
XX
PS Claim 9; SEQ ID NO 634; 598pp; Japanese.
XX
CC This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene,
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakaraensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakaraensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 326 AA;

Query Match 100.0%; Score 37; DB 8; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 138 HHPHG 142

RESULT 35
ABB66932
ID ABB66932 standard; protein; 327 AA.
XX
AC ABB66932;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 27588.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX

PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL11035.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 27588; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 327 AA;

Query Match 100.0%; Score 37; DB 4; Length 327;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 143 HHPHG 147

RESULT 36
ADS30838
ID ADS30838 standard; protein; 350 AA.
XX
AC ADS30838;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #19871.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
DR
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 19871; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 350 AA;

Query Match 100.0%; Score 37; DB 8; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 174 HHPHG 178

RESULT 37
ABB71820
ID ABB71820 standard; protein; 402 AA.
XX
AC ABB71820;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 42252.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX

PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR
DR N-PSDB; ABL15923.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 42252; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 402 AA;

Query Match 100.0%; Score 37; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 160 HHPHG 164

RESULT 38
ABB59479
ID ABB59479 standard; protein; 419 AA.
XX
AC ABB59479;
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 5229.
DE
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR
DR N-PSDB; ABL03582.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 5229; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 419 AA;

Query Match 100.0%; Score 37; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 253 HHPHG 257

RESULT 39
ABO77799
ID ABO77799 standard; protein; 466 AA.
XX
AC ABO77799;
XX

DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #9974.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX

PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD11370.
XX

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 26545; 455pp; English.
XX

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX
SQ Sequence 466 AA;

Query Match 100.0%; Score 37; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 135 HHPHG 139

RESULT 40
ABB79804
ID ABB79804 standard; protein; 525 AA.
XX

AC ABB79804;
XX
DT 25-NOV-2002 (first entry)
XX

DE Human histidine proline rich glycoprotein.
XX
KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cytosolic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 276..321
FT /note= "proline-rich domain"
FT Domain 350..497
FT /note= "Histidine-proline-rich domain, region also
FT specifically claimed in Claim 1"

XX WO200264621-A2.

PN 22-AUG-2002.

XX 14-FEB-2002; 2002WO-US004336.

XX 14-FEB-2001; 2001US-0268370P.

XX (ATTE-) ATTENUON LLC.

XX Donate F, Harris S, Plunkett ML, Mazar AP;

XX WPI; 2002-666989/71.

DR P-PSDB; ABN84910.

XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX Claim 1; Page 11; 82pp; English.

XX
CC The present sequence is the protein sequence of human histidine proline
CC rich glycoprotein (HPRG), a proteinaceous metal chelator that can be used
CC to inhibit angiogenesis and treat cancer. Claimed anti-angiogenic
CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)
CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see
CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,
CC endothelial cell proliferation or endothelial tube formation in vitro or
CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.
CC Also claimed are: chemically synthesised or recombinantly produced
CC peptide multimers; a diagnostically or therapeutically labeled anti-
CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
CC useful HPRG-related composition, comprising the diagnostically labeled
CC polypeptide, peptide or peptide multimer and a carrier; an antibody
CC specific for an epitope of HPRG that is present in the H/P domain of
CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,

CC or an antigen-binding fragment of the antibody; a method for inhibiting
CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
CC inducing apoptosis; a method for treating a subject having a disease or
CC condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis; a method of stimulating or inhibiting
CC angiogenesis in a subject; a method of detecting the presence of HPRG or
CC its cleavage product or its peptide in a biological sample; isolated
CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an
CC expression vector; transformed or transfected cells; a method of
CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of
CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;
CC an affinity ligand useful for binding to, or isolating, an HPRG-binding
CC molecule or cells expressing the binding molecule, comprising the
CC polypeptide, peptide or peptide multimer, immobilised to a solid support
CC or carrier; and a method of isolating HPRG-binding molecule, or isolating
CC or enriching cells expressing HPRG-binding site or receptor, from a
CC complex mixture. The compositions and methods are useful in diagnosing or
CC treating a disease or condition associated with undesired cell migration,
CC invasion, proliferation, or angiogenesis, such as cancer,
CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,
CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in
CC inhibiting the growth of primary tumours or metastases, and may also be
CC used in treating neurodegenerative diseases like Alzheimer's disease,
CC Parkinson's disease, and amyotrophic lateral sclerosis. The antibodies
CC are stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX

SQ Sequence 525 AA;

Query Match 100.0%; Score 37; DB 5; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

|||||

Db 384 HHPHG 388

Search completed: June 15, 2005, 14:15:48
Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-8
Perfect score: 37
Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	6	3	US-09-095-407-3
2	37	100.0	11	3	US-09-095-407-1
3	37	100.0	11	4	US-09-742-494-1
4	37	100.0	11	4	US-09-809-657-1
5	37	100.0	26	3	US-08-785-636-1
6	37	100.0	26	3	US-09-095-407-2
7	37	100.0	26	4	US-09-742-494-2
8	37	100.0	26	4	US-09-809-657-2
9	37	100.0	69	4	US-09-248-796A-24357
10	37	100.0	163	4	US-09-902-540-13395
11	37	100.0	220	4	US-09-270-767-61056
12	37	100.0	264	4	US-09-540-236-3168
13	37	100.0	273	4	US-09-949-016-6606
14	37	100.0	280	4	US-09-949-016-9101
15	37	100.0	466	4	US-09-252-991A-26545
16	37	100.0	525	4	US-09-976-594-64
17	37	100.0	525	4	US-09-919-039-62
18	37	100.0	539	4	US-09-518-959-8
19	37	100.0	539	4	US-09-518-959-9
20	37	100.0	618	4	US-09-252-991A-23373
21	37	100.0	810	4	US-09-248-796A-20281
22	37	100.0	923	4	US-09-270-767-45546
23	31	83.8	5	4	US-09-546-013-29
24	31	83.8	7	3	US-08-776-265-12
25	31	83.8	7	4	US-09-398-184-12
26	31	83.8	9	3	US-08-159-339A-1005
27	31	83.8	15	3	US-09-082-092-14

28	31	83.8	15	4	US-09-885-722A-14	Sequence 14, Appl
29	31	83.8	52	3	US-09-227-357-575	Sequence 575, App
30	31	83.8	69	4	US-09-910-009A-158	Sequence 158, App
31	31	83.8	75	4	US-09-716-129-73	Sequence 73, Appl
32	31	83.8	76	4	US-09-489-039A-10688	Sequence 10688, A
33	31	83.8	84	4	US-09-270-767-57094	Sequence 57094, A
34	31	83.8	91	4	US-09-270-767-36032	Sequence 36032, A
35	31	83.8	91	4	US-09-270-767-51249	Sequence 51249, A
36	31	83.8	92	4	US-09-248-796A-23534	Sequence 23534, A
37	31	83.8	99	4	US-09-640-211A-653	Sequence 653, App
38	31	83.8	102	4	US-09-270-767-58449	Sequence 58449, A
39	31	83.8	115	4	US-09-461-325-160	Sequence 160, App
40	31	83.8	115	4	US-10-012-542-160	Sequence 160, App
41	31	83.8	115	4	US-10-115-123-160	Sequence 160, App
42	31	83.8	125	4	US-09-248-796A-24231	Sequence 24231, A
43	31	83.8	127	4	US-09-430-221-2	Sequence 2, Appli
44	31	83.8	129	4	US-09-205-258-1135	Sequence 1135, Ap
45	31	83.8	132	1	US-08-470-179-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-09-095-407-3
; Sequence 3, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; Yip, Tai-Tung
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-095-407-3

Query Match 100.0%; Score 37; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
| | | | |
Db 2 HHPHG 6

RESULT 2
US-09-095-407-1
; Sequence 1, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; Yip, Tai-Tung
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,407
FILING DATE: 10-Jun-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,896
FILING DATE: 28-MAY-1993
APPLICATION NUMBER: WO PCT/US94/06064
FILING DATE: 27-MAY-1994
APPLICATION NUMBER: US 08/556,951
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 651-5325
TELEFAX: (713) 651-5246
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match 100.0%; Score 37; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
| | | | |
Db 2 HHPHG 6

RESULT 3
US-09-742-494-1
; Sequence 1, Application US/09742494
; Patent No. 6528320
; GENERAL INFORMATION:

; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-1

Query Match 100.0%; Score 37; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
| | | | |
Db 2 HHPHG 6

RESULT 4
US-09-809-657-1
; Sequence 1, Application US/09809657
; Patent No. 6734022
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USF/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-1

Query Match 100.0%; Score 37; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
| | | | |
Db 2 HHPHG 6

RESULT 5
US-08-785-636-1
; Sequence 1, Application US/08785636
; Patent No. 6027942
; GENERAL INFORMATION:
; APPLICANT: Yip, Tai-Tung
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Method and Apparatus for Desorption and Ionization of
; Analytes
; FILE REFERENCE: Hutchens
; CURRENT APPLICATION NUMBER: US/08/785,636
; CURRENT FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/068,896
; EARLIER FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Unknown

;
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: peptide
US-08-785-636-1

Query Match 100.0%; Score 37; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
Db 2 HHPHG 6

RESULT 6
US-09-095-407-2
; Sequence 2, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-DI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-095-407-2

Query Match 100.0%; Score 37; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
Db 2 HHPHG 6

RESULT 7

US-09-742-494-2
; Sequence 2, Application US/09742494
; Patent No. 6528320
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-2

Query Match 100.0%; Score 37; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
Db 2 HHPHG 6

RESULT 8
US-09-809-657-2
; Sequence 2, Application US/09809657
; Patent No. 6734022
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USF/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-2

Query Match 100.0%; Score 37; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
Db 2 HHPHG 6

RESULT 9
US-09-248-796A-24357
; Sequence 24357, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208

```

; SEQ ID NO 24357
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24357

Query Match      100.0%; Score 37; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      37 HHPHG 41

RESULT 10
US-09-902-540-13395
; Sequence 13395, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13395
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13395

Query Match      100.0%; Score 37; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      42 HHPHG 46

RESULT 11
US-09-270-767-61056
; Sequence 61056, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61056
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61056

Query Match      100.0%; Score 37; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      188 HHPHG 192

US-09-270-767-61056
```

```

RESULT 12
US-09-540-236-3168
; Sequence 3168, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3168
; LENGTH: 264
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3168

Query Match      100.0%; Score 37; DB 4; Length 264;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      230 HHPHG 234

RESULT 13
US-09-949-016-6606
; Sequence 6606, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6606
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6606

Query Match      100.0%; Score 37; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      229 HHPHG 233

RESULT 14
US-09-949-016-9101
; Sequence 9101, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9101
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9101

Query Match      100.0%; Score 37; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      236 HHPHG 240

RESULT 15
US-09-252-991A-26545
; Sequence 26545, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26545
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26545

Query Match      100.0%; Score 37; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      135 HHPHG 139

RESULT 16
US-09-976-594-64
; Sequence 64, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 085596CD1
US-09-976-594-64

Query Match      100.0%; Score 37; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      384 HHPHG 388

RESULT 17
US-09-919-039-62
; Sequence 62, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 085596CD1
US-09-919-039-62

Query Match      100.0%; Score 37; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      384 HHPHG 388

RESULT 18
US-09-518-959-8
; Sequence 8, Application US/09518959
; Patent No. 6548270
; GENERAL INFORMATION:
; APPLICANT: Dublin, Adrienne E
; APPLICANT: Erlander, Mark G
; APPLICANT: Huvar, Rene
; APPLICANT: Pyati, Jayashree
; TITLE OF INVENTION: DNA encoding human acid-sensing ion
; FILE REFERENCE: ORT-1197
; CURRENT APPLICATION NUMBER: US/09/518,959
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-959-8

Query Match      100.0%; Score 37; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||

```

Db 524 HHPHG 528

RESULT 19

US-09-518-959-9

; Sequence 9, Application US/09518959

; Patent No. 6548270

; GENERAL INFORMATION:

; APPLICANT: Dubin, Adrienne E

; APPLICANT: Erlander, Mark G

; APPLICANT: Huvar, Rene

; APPLICANT: Pyati, Jayashree

; TITLE OF INVENTION: DNA encoding human acid-sensing ion

; FILE REFERENCE: ORT-1197

; CURRENT APPLICATION NUMBER: US/09/518,959

; CURRENT FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 539

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-518-959-9

Query Match 100.0%; Score 37; DB 4; Length 539;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

Db 524 HHPHG 528

RESULT 20

US-09-252-991A-23373

; Sequence 23373, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23373

; LENGTH: 618

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23373

Query Match 100.0%; Score 37; DB 4; Length 618;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

Db 476 HHPHG 480

RESULT 21

US-09-248-796A-20281

; Sequence 20281, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 20281

; LENGTH: 810

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-20281

Query Match 100.0%; Score 37; DB 4; Length 810;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

Db 50 HHPHG 54

RESULT 22

US-09-270-767-45546

; Sequence 45546, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45546

; LENGTH: 923

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-45546

Query Match 100.0%; Score 37; DB 4; Length 923;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

Db 188 HHPHG 192

RESULT 23

US-09-546-013-29

; Sequence 29, Application US/09546013

; Patent No. 6610504

; GENERAL INFORMATION:

; APPLICANT: Yuan, Chong-Shen

; TITLE OF INVENTION: METHODS FOR ASSAYING S-ADENOSYLMETHIONINE-DEPENDENT Methyltransferase

; FILE REFERENCE: 10937-1652

; CURRENT APPLICATION NUMBER: US/09/546,013

; CURRENT FILING DATE: 2000-04-10

; EARLIER APPLICATION NUMBER: 09/347,878

; EARLIER FILING DATE: 1999-07-06

; EARLIER APPLICATION NUMBER: 09/457,205

; EARLIER FILING DATE: 1999-12-06

; NUMBER OF SEQ ID NOS: 94

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 29

; LENGTH: 5

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: peptide representing metal-binding protein surface
OTHER INFORMATION: domains from human histidine rich glycoprotein
OTHER INFORMATION: (HRG).
PUBLICATION INFORMATION:
JOURNAL: J. Chromatogr.
VOLUME: 604
ISSUE: 1
PAGES: 125-132
DATE: 1992
US-09-546-013-29

Query Match 83.8%; Score 31; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
|||||
Db 2 HHPH 5

RESULT 24
US-08-776-265-12
Sequence 12, Application US/08776265
Patent No. 6001631
GENERAL INFORMATION:
APPLICANT: BLANCHE, Francis
APPLICANT: CAMERON, Beatrice
APPLICANT: CROUZET, Joel
APPLICANT: FAMECHON, Alain
APPLICANT: FERRERO, Lucia
TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I. Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,265
FILING DATE: 24-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 03806.0394-00000
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4444
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

Query Match 83.8%; Score 31; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
:||||

Db 1 YHPHG 5
RESULT 25
US-09-398-184-12
Sequence 12, Application US/09398184
Patent No. 6649394
GENERAL INFORMATION:
APPLICANT: BLANCHE, Francis
APPLICANT: CAMERON, Beatrice
APPLICANT: CROUZET, Joel
APPLICANT: FAMECHON, Alain
APPLICANT: FERRERO, Lucia
TITLE OF INVENTION: No. 6649394el Topoisomerase IV, Corresponding
TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I. Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/398,184
FILING DATE: 17-Sep-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/776,265
FILING DATE: 24-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 03806.0394-00000
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4444
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-398-184-12

Query Match 83.8%; Score 31; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
:||||
Db 1 YHPHG 5

RESULT 26
US-08-159-339A-1005
Sequence 1005, Application US/08159339A
Patent No. 6037135
GENERAL INFORMATION:
APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Sette, Alessandro
APPLICANT: Celis, Esteban
TITLE OF INVENTION: HLA Binding peptides and Their
TITLE OF INVENTION: Uses

```

; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 1005:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-159-339A-1005

Query Match 83.8%; Score 31; DB 3; L
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0;

Qy 1 HHPH 4
    ||||
Db 3 HHPH 6

RESULT 27
US-09-082-092-14
; Sequence 14, Application US/09082092
; Patent No. 6251628
; GENERAL INFORMATION:
; APPLICANT: Nakao, Atsuhito
; APPLICANT: Moren, Anita
; APPLICANT: Heuchel, Rainer
; APPLICANT: Itoh, Susumu
; APPLICANT: Afrakhte, Mozghan
; APPLICANT: Souchelnytskyi, Serhiy
; APPLICANT: Brodin, Greger
; APPLICANT: Landstrom, Marene
; APPLICANT: Heldin, Nils-Erik
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: ten Dijke, Peter
; TITLE OF INVENTION: SMAD7 AND USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA

```

```

; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,092
; FILING DATE: 20-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,221
; FILING DATE: 20-MAY-1997
; APPLICATION NUMBER: 60/060,465
; FILING DATE: 30-SEP-1997
; APPLICATION NUMBER: 60/075,940
; FILING DATE: 25-FEB-1998
; APPLICATION NUMBER: 60/077,033
; FILING DATE: 06-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-09-082-092-14

Query Match      83.8%; Score 31; DB 3
Best Local Similarity 100.0%; Pred. No. 13;
Matches 4; Conservative 0; Mismatches

QY      1 HHPH 4
        ||||
Db      10 HHPH 13

RESULT 28
US-09-885-722A-14
; Sequence 14, Application US/09885722A
; Patent No. 6605443
; GENERAL INFORMATION:
; APPLICANT: Nakao, Atsuhito
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: ten Dijke, Peter
; TITLE OF INVENTION: SMAD7 AND USES THEREOF
; FILE REFERENCE: L00461.70117.US
; CURRENT APPLICATION NUMBER: US/09/885,722A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/082,092
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 60/077,033
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 60/075,940
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: US 60/060,465
; PRIOR FILING DATE: 1997-09-30
; PRIOR APPLICATION NUMBER: US 60/047,221
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 15
; TYPE: PRT

```



```

; ORGANISM: Mus musculus
US-09-885-722A-14

Query Match      83.8%; Score 31; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPH 4
      ||||
Db     10 HHPH 13

RESULT 29
US-09-227-357-575
; Sequence 575, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
```

```

; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 575
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-575

Query Match      83.8%; Score 31; DB 3; Length 52;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPH 4
      ||||
Db     30 HHPH 33

RESULT 30
US-09-910-009A-158
; Sequence 158, Application US/09910009A
; Patent No. 6727226
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Jacobsen, Richard
; APPLICANT: Jones, Robert M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Shen, Greg S.
; APPLICANT: Wagstaff, John D.
; TITLE OF INVENTION: Mu-Conopeptides
; FILE REFERENCE: 2314-242
; CURRENT APPLICATION NUMBER: US/09/910,009A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 60/219,619
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/245,157
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/264,319
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 60/277,270
; NUMBER OF SEQ ID NOS: 520
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 158
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Conus marmoreus
US-09-910-009A-158

Query Match      83.8%; Score 31; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 HHPH 4
Db 44 HHPH 47

RESULT 31
US-09-716-129-73
; Sequence 73, Application US/09716129
; Patent No. 6632920
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: PZ025P1
; CURRENT APPLICATION NUMBER: US/09/716,129
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/076,053
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,057
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,052
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,054
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,051
; PRIOR FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals stop translation
US-09-716-129-73

Query Match 83.8%; Score 31; DB 4; Length 75;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 23 HHPH 26

RESULT 32
US-09-489-039A-10688
; Sequence 10688, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10688
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10688

Query Match 83.8%; Score 31; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 44 HHPH 47

Db 72 HHPH 75

RESULT 33
US-09-270-767-57094
; Sequence 57094, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57094
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57094

Query Match 83.8%; Score 31; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 8 HHPH 11

RESULT 34
US-09-270-767-36032
; Sequence 36032, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36032
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-36032

Query Match 83.8%; Score 31; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 3 HHPH 6

RESULT 35
US-09-270-767-51249
; Sequence 51249, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51249
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster

US-09-270-767-51249

Query Match 83.8%; Score 31; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 3 HHPH 6

RESULT 36

US-09-248-796A-23534
; Sequence 23534, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 23534

; LENGTH: 92

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-23534

Query Match 83.8%; Score 31; DB 4; Length 92;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 73 HHPYG 77

RESULT 37

US-09-640-211A-653

; Sequence 653, Application US/09640211A

; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 653

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Eucalyptus grandis

US-09-640-211A-653

Query Match 83.8%; Score 31; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 9 HHPH 12

RESULT 38

US-09-270-767-58449

; Sequence 58449, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 58449

; LENGTH: 102

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-58449

Query Match 83.8%; Score 31; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 87 HHPH 90

RESULT 39

US-09-461-325-160

; Sequence 160, Application US/09461325A

; Patent No. 6475753

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: 94 Human Secreted Proteins

; FILE REFERENCE: PZ029P1

; CURRENT APPLICATION NUMBER: US/09/461,325A

; CURRENT FILING DATE: 1999-12-14

; EARLIER APPLICATION NUMBER: PCT/US99/13418

; EARLIER FILING DATE: 1999-06-15

; EARLIER APPLICATION NUMBER: 60/089,507

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,508

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,509

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,510

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/090,112

; EARLIER FILING DATE: 1998-06-22

; EARLIER APPLICATION NUMBER: 60/090,113

; EARLIER FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 532

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 160

; LENGTH: 115

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (96)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (100)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-461-325-160

Query Match 83.8%; Score 31; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
 ||||
Db 104 HHPH 107

RESULT 40
US-10-012-542-160
; Sequence 160, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-012-542-160

Query Match 83.8%; Score 31; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
 ||||
Db 104 HHPH 107

Search completed: June 15, 2005, 14:24:19
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-8
Perfect score: 37
Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	5	14 US-10-074-225A-8	Sequence 8, Appli
2	37	100.0	6	14 US-10-153-312A-1	Sequence 1, Appli
3	37	100.0	11	9 US-09-809-657-1	Sequence 1, Appli
4	37	100.0	11	9 US-09-848-512A-1	Sequence 1, Appli
5	37	100.0	11	9 US-09-742-494-1	Sequence 1, Appli
6	37	100.0	11	14 US-10-153-312A-2	Sequence 2, Appli
7	37	100.0	16	14 US-10-153-312A-3	Sequence 3, Appli
8	37	100.0	26	9 US-09-809-657-2	Sequence 2, Appli
9	37	100.0	26	9 US-09-848-512A-2	Sequence 2, Appli
10	37	100.0	26	9 US-09-742-494-2	Sequence 2, Appli
11	37	100.0	26	14 US-10-153-312A-4	Sequence 4, Appli

12	37	100.0	84	15 US-10-424-599-277019	Sequence 277019,
13	37	100.0	97	14 US-10-106-698-4437	Sequence 4437, Ap
14	37	100.0	101	14 US-10-074-225A-6	Sequence 6, Appli
15	37	100.0	116	16 US-10-437-963-149790	Sequence 149790,
16	37	100.0	129	16 US-10-425-115-324288	Sequence 324288,
17	37	100.0	138	16 US-10-437-963-142824	Sequence 142824,
18	37	100.0	148	14 US-10-074-225A-5	Sequence 5, Appli
19	37	100.0	151	9 US-09-864-761-42051	Sequence 42051, A
20	37	100.0	152	16 US-10-437-963-128472	Sequence 128472,
21	37	100.0	185	15 US-10-425-114-58759	Sequence 58759, A
22	37	100.0	259	15 US-10-282-122A-62936	Sequence 62936, A
23	37	100.0	273	10 US-09-981-151A-64	Sequence 64, Appl
24	37	100.0	273	10 US-09-981-151A-65	Sequence 65, Appl
25	37	100.0	273	15 US-10-295-027-344	Sequence 344, App
26	37	100.0	273	15 US-10-295-027-1250	Sequence 1250, Ap
27	37	100.0	273	15 US-10-072-012-593	Sequence 593, App
28	37	100.0	273	15 US-10-072-012-596	Sequence 596, App
29	37	100.0	273	15 US-10-188-832-185	Sequence 185, App
30	37	100.0	350	15 US-10-369-493-19871	Sequence 19871, A
31	37	100.0	460	15 US-10-424-599-205101	Sequence 205101,
32	37	100.0	492	16 US-10-808-807-8	Sequence 8, Appli
33	37	100.0	525	10 US-09-919-039-62	Sequence 62, Appl
34	37	100.0	525	14 US-10-074-225A-1	Sequence 1, Appli
35	37	100.0	525	17 US-10-868-577A-42	Sequence 42, Appl
36	37	100.0	526	14 US-10-074-225A-3	Sequence 3, Appli
37	37	100.0	539	10 US-09-772-180A-2	Sequence 2, Appli
38	37	100.0	539	10 US-09-772-180A-8	Sequence 8, Appli
39	37	100.0	539	15 US-10-295-027-290	Sequence 290, App
40	37	100.0	587	10 US-09-772-180A-4	Sequence 4, Appli
41	37	100.0	19695	15 US-10-084-846A-3	Sequence 3, Appli
42	31	83.8	5	14 US-10-419-802-1	Sequence 1, Appli
43	31	83.8	5	15 US-10-373-238-25	Sequence 25, Appl
44	31	83.8	7	15 US-10-646-807-12	Sequence 12, Appl
45	31	83.8	10	17 US-10-883-760-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-10-074-225A-8
; Sequence 8, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-8

Query Match 100.0%; Score 37; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
| | | | |
Db 1 HHPHG 5

```

RESULT 2
US-10-153-312A-1
; Sequence 1, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
US-10-153-312A-1

Query Match      100.0%; Score 37; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
        |||||
Db      2 HHPHG 6

RESULT 3
US-09-809-657-1
; Sequence 1, Application US/09809657
; Patent No. US20010014479A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USF/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-1

Query Match      100.0%; Score 37; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
        |||||
Db      2 HHPHG 6

RESULT 4
US-09-848-512A-1
; Sequence 1, Application US/09848512A
; Patent No. US20020037517A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS FOR SEQUENCING BIOPOLYMERS
```

```

; FILE REFERENCE: P00798USG/09306611
; CURRENT APPLICATION NUMBER: US/09/848,512A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-848-512A-1

Query Match      100.0%; Score 37; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
        |||||
Db      2 HHPHG 6

RESULT 5
US-09-742-494-1
; Sequence 1, Application US/09742494
; Patent No. US20020155620A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-1

Query Match      100.0%; Score 37; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
        |||||
Db      2 HHPHG 6

RESULT 6
US-10-153-312A-2
; Sequence 2, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
```


US-10-153-312A-2

Query Match 100.0%; Score 37; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
Db 2 HHPHG 6

RESULT 7

US-10-153-312A-3
; Sequence 3, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide

US-10-153-312A-3

Query Match 100.0%; Score 37; DB 14; Length 16;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
Db 2 HHPHG 6

RESULT 8

US-09-809-657-2
; Sequence 2, Application US/09809657
; Patent No. US20010014479A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USF/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-809-657-2

Query Match 100.0%; Score 37; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
Db 2 HHPHG 6

RESULT 9
US-09-848-512A-2
; Sequence 2, Application US/09848512A
; Patent No. US20020037517A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS FOR SEQUENCING BIOPOLYMERS
; FILE REFERENCE: P00798USG/09306611
; CURRENT APPLICATION NUMBER: US/09/848,512A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: HOMO SAPIENS

US-09-848-512A-2

Query Match 100.0%; Score 37; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
Db 2 HHPHG 6

RESULT 10

US-09-742-494-2
; Sequence 2, Application US/09742494
; Patent No. US20020155620A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-742-494-2

Query Match 100.0%; Score 37; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|
Db 2 HHPHG 6

RESULT 11

US-10-153-312A-4
; Sequence 4, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24

```

; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
US-10-153-312A-4

Query Match      100.0%; Score 37; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
        |||||
Db      2 HHPHG 6

RESULT 12
US-10-424-599-277019
; Sequence 277019, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 277019
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92167C.1.pep
US-10-424-599-277019

Query Match      100.0%; Score 37; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
        |||||
Db      52 HHPHG 56

RESULT 13
US-10-106-698-4437
; Sequence 4437, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
```

```

; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4437
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4437

Query Match      100.0%; Score 37; DB 14; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
        |||||
Db      53 HHPHG 57

RESULT 14
US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match      100.0%; Score 37; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
        |||||
Db      13 HHPHG 17

RESULT 15
US-10-437-963-149790
; Sequence 149790, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149790
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50089C.1.pap
US-10-437-963-149790

Query Match 100.0%; Score 37; DB 16; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 89 HHPHG 93

RESULT 16
US-10-425-115-324288
; Sequence 324288, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 324288
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58824C.1.pap
US-10-425-115-324288

Query Match 100.0%; Score 37; DB 16; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 85 HHPHG 89

RESULT 17
US-10-437-963-142824
; Sequence 142824, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142824
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43793C.1.pap
US-10-437-963-142824

Query Match 100.0%; Score 37; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 36 HHPHG 40

RESULT 18
US-10-074-225A-5
; Sequence 5, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-5

Query Match 100.0%; Score 37; DB 14; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 35 HHPHG 39

RESULT 19
US-09-864-761-42051
; Sequence 42051, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04

```
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42051
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL122010.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EST HUMAN HIT: BE182360.1, EVALUE 3.00e-64
; OTHER INFORMATION: SWISSPROT HIT: O95377, EVALUE 2.00e-89
US-09-864-761-42051

Query Match          100.0%; Score 37; DB 9; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      128 HHPHG 132

RESULT 20
US-10-437-963-128472
; Sequence 128472, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 128472
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30821C.1.pep
US-10-437-963-128472

Query Match          100.0%; Score 37; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      24 HHPHG 28

RESULT 21
US-10-425-114-58759
; Sequence 58759, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58759
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700214339_FLI.pep
US-10-425-114-58759

Query Match          100.0%; Score 37; DB 15; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      117 HHPHG 121

RESULT 22
US-10-282-122A-62936
; Sequence 62936, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
```

```
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62936
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-282-122A-62936

Query Match 100.0%; Score 37; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 229 HHPHG 233

RESULT 23
US-09-981-151A-64
; Sequence 64, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shimkets, Richard A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
```

```
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-151A-64

Query Match 100.0%; Score 37; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 229 HHPHG 233

RESULT 24
US-09-981-151A-65
; Sequence 65, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shimkets, Richard A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
```

```
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (41)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-09-981-151A-65

Query Match      100.0%; Score 37; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      229 HHPHG 233

RESULT 25
US-10-295-027-344
; Sequence 344, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
```

```
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 344
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-344

Query Match      100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
      |||||
Db      229 HHPHG 233

RESULT 26
US-10-295-027-1250
; Sequence 1250, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1250
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1250

Query Match      100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 1 HHPHG 5
|||||
Db 229 HHPHG 233

RESULT 27

US-10-072-012-593
; Sequence 593, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 593
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-593

Query Match 100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
|||||

Db 229 HHPHG 233

RESULT 28

US-10-072-012-596
; Sequence 596, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 596
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (41)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-072-012-596

Query Match 100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
|||||

Db 229 HHPHG 233
|||||
US-10-188-832-185
; Sequence 185, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-185
Query Match 100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHPHG 5
|||||
Db 229 HHPHG 233
RESULT 30
US-10-369-493-19871
; Sequence 19871, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19871
; LENGTH: 350
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-19871
Query Match 100.0%; Score 37; DB 15; Length 350;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 174 HHPHG 178
RESULT 31
US-10-424-599-205101
; Sequence 205101, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 205101
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(460)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_27235C.1.pep
US-10-424-599-205101
Query Match 100.0%; Score 37; DB 15; Length 460;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
|||||
Db 15 HHPHG 19
RESULT 32
US-10-808-807-8
; Sequence 8, Application US/10808807
; Publication No. US20040253663A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont deNemours and Co., Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2365 US NA
; CURRENT APPLICATION NUMBER: US/10/808,807
; CURRENT FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/477,874
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Pantoea agglomerans strain DC404
US-10-808-807-8
Query Match 100.0%; Score 37; DB 16; Length 492;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
|||||
Db 286 HHPHG 290

RESULT 33
US-09-919-039-62
; Sequence 62, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 085596CD1
US-09-919-039-62

Query Match 100.0%; Score 37; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 384 HHPHG 388

RESULT 34
US-10-074-225A-1
; Sequence 1, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-1

Query Match 100.0%; Score 37; DB 14; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 384 HHPHG 388

RESULT 35
US-10-868-577A-42
; Sequence 42, Application US/10868577A
; Publication No. US20050032697A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al.

; TITLE OF INVENTION: HEPARIN BINDING VEGFR-3 LIGANDS
; FILE REFERENCE: 28967/39359A
; CURRENT APPLICATION NUMBER: US/10/868,577A
; CURRENT FILING DATE: 2004-06-14
; PRIOR APPLICATION NUMBER: US 60/478,390
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 10/669,176
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-868-577A-42

Query Match 100.0%; Score 37; DB 17; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 384 HHPHG 388

RESULT 36
US-10-074-225A-3
; Sequence 3, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match 100.0%; Score 37; DB 14; Length 526;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 333 HHPHG 337

RESULT 37
US-09-772-180A-2
; Sequence 2, Application US/09772180A
; Publication No. US20030027749A1
; GENERAL INFORMATION:
; APPLICANT: David C. Harrison
; APPLICANT: John Davis
; APPLICANT: Sharon Bingham
; APPLICANT: Trudy R. Doe
; APPLICANT: Simon Topp
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30021-C1
; CURRENT APPLICATION NUMBER: US/09/772,180A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 09/063,848

```
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 9708936.1
; PRIOR FILING DATE: 1997-05-01
; PRIOR APPLICATION NUMBER: 97310289.0
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 9803566.0
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 539
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-772-180A-2

Query Match      100.0%; Score 37; DB 10; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
      |||||
Db      524 HHPHG 528

RESULT 38
US-09-772-180A-8
; Sequence 8, Application US/09772180A
; Publication No. US20030027749A1
; GENERAL INFORMATION:
; APPLICANT: David C. Harrison
; APPLICANT: John Davis
; APPLICANT: Sharon Bingham
; APPLICANT: Trudy R. Doe
; APPLICANT: Simon Topp
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30021-C1
; CURRENT APPLICATION NUMBER: US/09/772,180A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 09/063,848
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 9708936.1
; PRIOR FILING DATE: 1997-05-01
; PRIOR APPLICATION NUMBER: 97310289.0
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 9803566.0
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 539
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-772-180A-8

Query Match      100.0%; Score 37; DB 10; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
      |||||
Db      524 HHPHG 528

RESULT 39
US-10-295-027-290
; Sequence 290, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynne, Richard
```

```
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 290
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-290

Query Match      100.0%; Score 37; DB 15; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HHPHG 5
      |||||
Db      524 HHPHG 528

RESULT 40
US-09-772-180A-4
; Sequence 4, Application US/09772180A
; Publication No. US20030027749A1
; GENERAL INFORMATION:
; APPLICANT: David C. Harrison
; APPLICANT: John Davis
; APPLICANT: Sharon Bingham
; APPLICANT: Trudy R. Doe
; APPLICANT: Simon Topp
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30021-C1
; CURRENT APPLICATION NUMBER: US/09/772,180A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 09/063,848
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 9708936.1
; PRIOR FILING DATE: 1997-05-01
; PRIOR APPLICATION NUMBER: 97310289.0
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 9803566.0
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
```

```
;
; LENGTH: 587
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE: FEATURE
; NAME/KEY: UNSURE
; LOCATION: (328) (356) (357) (358) (359) (360) (361) (362) (363) (364) (365) (366)
; LOCATION: (367) (368) (369) (370) (371) (372) (373) (374) (375) (376) (377) (378)
; LOCATION: (379) (380) (381) (382) (383) (384) (385) (386) (387) (388) (389) (390)
; LOCATION: (391) (392) (393) (394) (395) (396) (397) (398) (399) (400) (401) (402)
; LOCATION: (403) (404) (405) (406) (407) (408) (409) (410) (411) (412) (413) (414)
; LOCATION: (415) (416) (417)
; OTHER INFORMATION: OTHER INFORMATION: Partial Amino Acid Sequence
US-09-772-180A-4

Query Match      100.0%; Score 37; DB 10; Length 587;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1 HHPHG 5
      |||||
Db      572 HHPHG 576

Search completed: June 15, 2005, 14:50:18
Job time : 80.75 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-8
Perfect score: 37
Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	37	100.0	419	1	S70617	homeotic protein g
2	37	100.0	445	2	A60488	histidine-rich gly
3	37	100.0	491	2	H83979	lysine decarboxyla
4	37	100.0	525	1	KGHUGH	histidine-rich gly
5	37	100.0	602	2	D75618	hypothetical prote
6	37	100.0	721	2	E70766	hypothetical prote
7	37	100.0	753	1	D72660	probable aldehyde
8	37	100.0	1441	2	T13717	CRAG protein - fru
9	31	83.8	76	2	F97793	hypothetical prote
10	31	83.8	84	2	AH3162	IS21 family transp
11	31	83.8	94	2	A49832	DNA gyrase A - Sta
12	31	83.8	95	2	S49447	DNA topoisomerase
13	31	83.8	95	2	S54254	DNA topoisomerase
14	31	83.8	105	2	S62940	probable membrane
15	31	83.8	110	2	F69760	hypothetical prote
16	31	83.8	113	2	E82687	cytochrome O ubiqu
17	31	83.8	130	2	S14983	extensin class I (
18	31	83.8	134	2	D72634	hypothetical prote
19	31	83.8	136	2	E72759	hypothetical prote
20	31	83.8	146	2	I40541	hypothetical prote
21	31	83.8	147	2	I39542	DNA gyrase - Aerom
22	31	83.8	147	2	AC1898	urease accessory p
23	31	83.8	148	2	F89956	conserved hypothet
24	31	83.8	155	2	S25846	homeotic protein H
25	31	83.8	156	2	T31664	DAP-kinase homolog
26	31	83.8	164	2	I58201	MHC class I antige
27	31	83.8	171	2	T13554	hypothetical prote
28	31	83.8	178	2	T36013	probable integral
29	31	83.8	188	2	B49773	ecdysone-dependent

30	31	83.8	193	2	E83864	thioredoxin BH1717
31	31	83.8	211	2	S50686	hypothetical prote
32	31	83.8	213	2	G87457	guanylate kinase l
33	31	83.8	215	2	JX0244	pyroglutamyl-pepti
34	31	83.8	217	2	T06455	Myb26 protein - ga
35	31	83.8	223	2	I59173	glutamate decarbox
36	31	83.8	227	2	S35735	DNA gyrase chain A
37	31	83.8	228	2	T40889	hypothetical prote
38	31	83.8	228	2	S65426	pyruvate decarboxy
39	31	83.8	232	2	H82360	conserved hypothet
40	31	83.8	236	2	AB0259	probable phage min
41	31	83.8	244	2	A84729	hypothetical prote
42	31	83.8	245	2	D70882	probable dapB prot
43	31	83.8	251	2	AB0957	conserved hypothet
44	31	83.8	252	2	AE3631	nitrous-oxide redu
45	31	83.8	253	2	F84845	hypothetical prote

ALIGNMENTS

RESULT 1

S70617 homeotic protein goosecoid - fruit fly (Drosophila melanogaster)

N;Alternate names: homeobox protein goosecoid

C;Species: Drosophila melanogaster

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: S70617

R;Hahn, M.; Jaeckle, H.

EMBO J. 15, 3077-3084, 1996

A;Title: Drosophila goosecoid participates in neural development but not in body axis for

A;Reference number: S70617; MUID:96272167; PMID:8670808

A;Accession: S70617

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-419 <HAH>

A;Cross-references: UNIPROT:P54366; EMBL:U52968; NID:g1399586; PIDN:AAB17948.1; PID:g1399

C;Genetics:

A;Gene: gsc

A;Cross-references: FlyBase:FBgn0010323

A;Map position: 2

C;Function:

A;Description: plays a role in neurogenesis in post-gastrula Drosophila embryos

A;Note: not required for gastrulation like Xenopus goosecoid; expressed most strongly in

C;Superfamily: fruit fly homeotic protein goosecoid; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;287-343/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 37; DB 1; Length 419;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 253 HHPHG 257

RESULT 2

A60488

histidine-rich glycoprotein - bovine (fragments)

N;Alternate names: autorosette inhibition factor

C;Species: Bos primigenius taurus (cattle)

C;Date: 19-Mar-1993 #sequence_revision 23-Mar-1995 #text_change 07-Jul-1995

C;Accession: S35687; JC2196; A60488

R;Sorensen, C.B.; Krogh-Pedersen, H.; Petersen, T.E.

FEBS Lett. 328, 285-290, 1993

A;Title: Determination of the disulphide bridge arrangement of bovine histidine-rich glyc

A;Reference number: S35687; MUID:93351678; PMID:8348977

A;Accession: S35687

A;Molecule type: protein

A;Residues: 1-25;26-52;57-75;82-88;95-119;146-173;175-206;210-309;313-445 <SOR>

A;Note: 355-Gln and 368-Tyr were also found

R;Halkier, T.; Andersen, H.; Vestergaard, A.; Magnusson, S.

Db	 411 HHPHG 415	
RESULT 6		
E70766	hypothetical protein Rv2082 - Mycobacterium tuberculosis (strain H37RV) C:Species: Mycobacterium tuberculosis C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004 C:Accession: E70766 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; MUID:98295987; PMID:9634230 A:Accession: E70766 A:Status: preliminary; nucleic acid sequence not shown; translation not shown A:Molecule type: DNA A:Residues: 1-721 <COL> A:Cross-references: UNIPROT:Q10690; GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98194. A:Experimental source: strain H37RV C:Genetics: A:Gene: Rv2082	
	Query Match 100.0%; Score 37; DB 2; Length 721; Best Local Similarity 100.0%; Pred. No. 61; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 	
Db	591 HHPHG 595	
RESULT 7		
D72660	probable aldehyde oxidoreductase APE0708 - Aeropyrum pernix (strain K1) C:Species: Aeropyrum pernix C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004 C:Accession: D72660 R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum A:Reference number: A72450; MUID:99310339; PMID:10382966 A:Accession: D72660 A:Molecule type: DNA A:Residues: 1-753 <KAW> A:Cross-references: UNIPROT:Q9YE62; DDBJ:AP000060; NID:g5104188; PIDN:BAA79684.1; PID:g5 A:Experimental source: strain K1 C:Genetics: A:Gene: APE0708 C:Superfamily: carbon monoxide dehydrogenase molybdoprotein	
	Query Match 100.0%; Score 37; DB 1; Length 753; Best Local Similarity 100.0%; Pred. No. 64; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 	
Db	696 HHPHG 700	
RESULT 8		
T13717	CRAG protein - fruit fly (Drosophila melanogaster) C:Species: Drosophila melanogaster C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004 C:Accession: T13717 R;Xu, X.Z.; Wes, P.D.; Chen, H.; Li, H.S.; Yu, M.; Morgan, S.; Liu, Y.; Montell, C. J. Biol. Chem. 273, 31297, 1998 A:Title: Retinal targets for calmodulin include proteins implicated in synaptic transmissi	

	A;Reference number: Z17709; MUID:99030403; PMID:9813038 A;Accession: T13717 A;Status: preliminary; translated from GB/EMBL/DDBJ A:Molecule type: mRNA A;Residues: 1-1441 <XUX> A;Cross-references: UNIPROT:Q96957; EMBL:Y17918; NID:g3893100; PIDN:CAA76938.1; PID:g389; C:Genetics: A:Gene: CRAG A:Cross-references: FlyBase:FBgn0025864	
	Query Match 100.0%; Score 37; DB 2; Length 1441; Best Local Similarity 100.0%; Pred. No. 1.2e+02; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 	
Db	1119 HHPHG 1123	
RESULT 9		
F97793	hypothetical protein RC0750 [imported] - Rickettsia conorii (strain Malish 7) C:Species: Rickettsia conorii C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004 C:Accession: F97793 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; RoScience 293, 2093-2098, 2001 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii. A:Reference number: A97700; MUID:21442074; PMID:11557893 A:Accession: F97793 A:Status: preliminary A:Molecule type: DNA A:Residues: 1-76 <KUR> A:Cross-references: UNIPROT:Q92HM1; GB:AE006914; PIDN:AAL03288.1; PID:g15619845; GSPDB:GC C:Genetics: A:Gene: RC0750	
	Query Match 83.8%; Score 31; DB 2; Length 76; Best Local Similarity 80.0%; Pred. No. 65; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 :	
Db	7 HHPYG 11	
RESULT 10		
AH3162	IS21 family transposase (truncated) tnp [imported] - Agrobacterium tumefaciens (strain C C:Species: Agrobacterium tumefaciens C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004 C:Accession: AH3162 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L. erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyaivin, T.; Levy, R.; Li, M.; McClell; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W. A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. A:Reference number: AB2577; MUID:21608550; PMID:11743193 A:Accession: AH3162 A:Status: preliminary A:Molecule type: DNA A:Residues: 1-84 <KUR> A:Cross-references: UNIPROT:Q8UKS4; GB:AE008687; PIDN:AAL45718.1; PID:gl7743448; GSPDB:GA A:Experimental source: strain C58 (Dupont) C:Genetics: A:Gene: tnp A:Genome: plasmid	
	Query Match 83.8%; Score 31; DB 2; Length 84; Best Local Similarity 100.0%; Pred. No. 72; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

Qy	1 HHPH 4 :	
Db	47 HHPH 50	
RESULT 11		
A49832		DNA gyrase A - Staphylococcus epidermidis (fragment)
C;Species:	Staphylococcus epidermidis	
C;Date:	07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004	
C;Accession:	A49832	
R;Sreedharan, S.;	Peterson, L.R.; Fisher, L.M.	
Antimicrob. Agents Chemother.	35, 2151-2154, 1991	
A;Title:	Ciprofloxacin resistance in coagulase-positive and -negative staphylococci: role of	
.		
A;Reference number:	A49832; MUID:92102204; PMID:1662027	
A;Accession:	A49832	
A;Status:	preliminary	
A;Molecule type:	DNA	
A;Residues:	1-94 <SRE>	
A;Cross-references:	UNIPROT:P54112; GB:S72603; NID:g240996; PIDN:AAB20672.1; PID:g240997	
A;Note:	sequence extracted from NCBI backbone (NCBIN:72603, NCBIP:72604)	
C;Superfamily:	DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (
C;Keywords:	ATP	
F;1-94/Domain:	phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragm	
Query Match	83.8%; Score 31; DB 2; Length 94;	
Best Local Similarity	80.0%; Pred. No. 81;	
Matches	4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 :	
Db	78 YHPHG 82	
RESULT 12		
S49447		DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain A - Acinetobacter baumannii (fra
N;Alternate names:	DNA gyrase chain A	
C;Species:	Acinetobacter baumannii	
C;Date:	20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004	
C;Accession:	S49447	
R;Vila, J.; Ruiz, J.;	Goni, P.; Marcos, M.; Jimenez de Anta, M.T.	
submitted to the EMBL Data Library,	October 1994	
A;Reference number:	S49447	
A;Accession:	S49447	
A;Molecule type:	DNA	
A;Residues:	1-95 <VIL>	
A;Cross-references:	UNIPROT:Q43906; EMBL:X82165; NID:g558546; PIDN:CAA57655.1; PID:g5585	
C;Genetics:		
A;Gene:	gyrA	
C;Superfamily:	DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (
C;Keywords:	ATP; DNA binding; DNA replication; isomerase	
F;1-95/Domain:	phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragm	
Query Match	83.8%; Score 31; DB 2; Length 95;	
Best Local Similarity	80.0%; Pred. No. 82;	
Matches	4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 :	
Db	8 YHPHG 12	
RESULT 13		
S54254		DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain A - Salmonella typhimurium (frag
N;Alternate names:	DNA gyrase chain A; type II DNA topoisomerase chain A	
C;Species:	Salmonella typhimurium	
C;Date:	08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004	
C;Accession:	S54254	
R;Ruiz, J.; Castro, D.;	Goni, P.; Borrego, J.J.; Vila, J.	

submitted to the EMBL Data Library,	April 1995	
A;Reference number:	S54254	
A;Accession:	S54254	
A;Molecule type:	DNA	
A;Residues:	1-95 <RUI>	
A;Cross-references:	UNIPROT:P37411; EMBL:X86695; NID:g799016; PIDN:CAA60388.1; PID:g79901	
C;Comment:	The A chain is the target of quinolone antibiotics such as nalidixic acid and	
C;Genetics:		
A;Gene:	gyrA	
C;Superfamily:	DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (
C;Keywords:	ATP; DNA binding; DNA replication; heterotetramer; isomerase	
F;1-95/Domain:	phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragme	
F;51/Active site:	Tyr #status Predicted	
Query Match	83.8%; Score 31; DB 2; Length 95;	
Best Local Similarity	80.0%; Pred. No. 82;	
Matches	4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPHG 5 :	
Db	6 YHPHG 10	
RESULT 14		
S62940		probable membrane protein YNL028w - yeast (Saccharomyces cerevisiae)
N;Alternate names:	hypothetical protein N2758	
C;Species:	Saccharomyces cerevisiae	
C;Date:	27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004	
C;Accession:	S62940; S62950	
R;Andre, B.;	Iraqi Houssaini, I.;	Urrestarazu, L.A.;
submitted to the Protein Sequence Database,	April 1996	
A;Reference number:	S62920	
A;Accession:	S62940	
A;Molecule type:	DNA	
A;Residues:	1-105 <AND>	
A;Cross-references:	UNIPROT:P53967; EMBL:Z71305; NID:gl301864; PID:gl301866; MIPS:YNL028v	
A;Experimental source:	strain S288C	
R;Duesterhoeft, A.;	Floeth, M.;	Fritz, C.;
Heuss-Neitzel, D.;	Hilbert, H.;	Moestl, D.
submitted to the Protein Sequence Database,	April 1996	
A;Reference number:	S62944	
A;Accession:	S62950	
A;Molecule type:	DNA	
A;Residues:	1-105 <DUE>	
A;Cross-references:	EMBL:Z71305; NID:gl301864; PID:gl301866; MIPS:YNL028w	
A;Experimental source:	strain S288C	
C;Genetics:		
A;Cross-references:	SGD:S0004973	
A;Map position:	14L	
C;Superfamily:	Saccharomyces probable membrane protein YNL028w	
C;Keywords:	transmembrane protein	
F;23-39/Domain:	transmembrane #status predicted <TMM>	
Query Match	83.8%; Score 31; DB 2; Length 105;	
Best Local Similarity	100.0%; Pred. No. 90;	
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 HHPH 4 :	
Db	85 HHPH 88	
RESULT 15		
F69760		hypothetical protein yckD - Bacillus subtilis
C;Species:	Bacillus subtilis	
C;Date:	05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004	
C;Accession:	F69760	
R;Kunst, F.;	Ogasawara, N.;	Moszer, I.;
Albertini, A.M.;	Alloni, G.;	Azevedo, V.;
Bertero, C.;	Bron, S.;	Brouillet, S.;
Bruschi, C.V.;	Caldwell, B.;	Capuano, V.;
Carter, N.M.;	Choi, A.;	Ehrlich, S.D.;
Emmerson, P.T.;	Entian, K.D.;	Errington, J.;
Fabret, C.;	Ferrari, E. Nature 390, 249-256, 1997	
A;Authors:	Foulger, D.;	Fritz, C.;
Fujita, M.;	Fujita, Y.;	Fuma, S.;
Galizzi, A.;	Galleri	

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadai, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69760
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-110 <KUN>
A;Cross-references: UNIPROT:P42402; GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CAB12134.
A;Experimental source: strain 168
C;Genetics:
A;Gene: yckD
C;Superfamily: *Bacillus subtilis* hypothetical protein yckD

Query Match 83.8%; Score 31; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 99 HHPH 102

RESULT 16
E82687
cytochrome O ubiquinol oxidase, subunit IV XF1387 [imported] - *Xylella fastidiosa* (strain
C;Species: *Xylella fastidiosa*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82687
R;anonymouse, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82687
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <SIM>
A;Cross-references: UNIPROT:Q9PDJ2; GB:AE003970; GB:AE003849; NID:g9106392; PIDN:AAF8419
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1387

Query Match 83.8%; Score 31; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 3 HHPH 6

RESULT 17
S14983

extensin class I (clone w10-1 L) - tomato (fragment)
C;Species: *Lycopersicon esculentum* (tomato)
C;Date: 07-May-1998 #sequence_revision 15-May-1998 #text_change 17-Jul-1998
C;Accession: S14983
R;Showalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.
Plant Mol. Biol. 16, 547-565, 1991
A;Title: Tomato extensin and extensin-like cDNAs: structure and expression in response to
A;Reference number: S14970; MUID:91329690; PMID:1714316
A;Accession: S14983
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-130 <SHO>
A;Cross-references: EMBL:X55694
A;Experimental source: cv. UC82B
C;Keywords: cell wall; glycoprotein; hydroxyproline

Query Match 83.8%; Score 31; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 58 HHPH 61

RESULT 18
D72634

hypothetical protein APE1531 - *Aeropyrum pernix* (strain K1)
C;Species: *Aeropyrum pernix*
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: D72634
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aeropyr*
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: D72634
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-134 <KAW>
A;Cross-references: UNIPROT:Q9YBR8; DBJ:AP000061; NID:g5104821; PIDN:BAA80530.1; PID:g51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1531
C;Superfamily: *Aeropyrum pernix* hypothetical protein APE1531

Query Match 83.8%; Score 31; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
Db 107 HHPH 110

RESULT 19
E72759

hypothetical protein APE0066 - *Aeropyrum pernix* (strain K1)
C;Species: *Aeropyrum pernix*
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E72759
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aeropyr*
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-136 <KAW>
A;Cross-references: UNIPROT:Q9YG35; DBJ:AP000058; NID:g5103388; PIDN:BAA78975.1; PID:dlc

A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0066
C;Superfamily: Aeropyrum pernix hypothetical protein APE0066

Query Match 83.8%; Score 31; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 132 HHPH 135

RESULT 20
I40541
hypothetical protein X - Pseudomonas solanacearum
C;Species: Pseudomonas solanacearum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C;Accession: I40541
R;Huang, J.; Carney, B.F.; Denny, T.P.; Weissinger, A.K.; Schell, M.A.
J. Bacteriol. 177, 1259-1267, 1995
A;Title: A complex network regulates expression of eps and other virulence genes of Pseudomonas
A;Reference number: I40539; MUID:95173103; PMID:7868600
A;Accession: I40541
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-146 <RES>
A;Cross-references: UNIPROT:Q45417; EMBL:U18135; NID:g603068; PIDN:AAA66241.1; PID:g603068

Query Match 83.8%; Score 31; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 142 HHPH 145

RESULT 21
I39542
DNA gyrase - Aeromonas salmonicida (fragment)
C;Species: Aeromonas salmonicida
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C;Accession: I39542
R;Oppegaard, H.; Sorum, H.
Antimicrob. Agents Chemother. 38, 2460-2464, 1994
A;Title: gyrA mutations in quinolone-resistant isolates of the fish pathogen Aeromonas salmonicida
A;Reference number: I39542; MUID:95142596; PMID:7840589
A;Accession: I39542
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-147 <RES>
A;Cross-references: UNIPROT:P48369; GB:L42453; NID:g832830; PIDN:AAA87239.1; PID:g832831
C;Genetics:
A;Gene: gyrA
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragment)
F;1-147/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragment)

Query Match 83.8%; Score 31; DB 2; Length 147;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
:|||
Db 45 YHPHG 49

RESULT 22
AC1898
urease accessory protein E [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 12-Jul-2004
C;Accession: AC1898
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC1898
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-147 <KUR>
A;Cross-references: UNIPROT:Q8YYW0; GB:BA000019; PIDN:BAB72690.1; PID:gl7130078; GSPDB:GN000001
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0733
C;Superfamily: urease accessory protein (nickel metallochaperone) UreE

Query Match 83.8%; Score 31; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 144 HHPH 147

RESULT 23
F89956
conserved hypothetical protein SA1543 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: F89956
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Hoshino, T.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: F89956
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-148 <KUR>
A;Cross-references: UNIPROT:Q99TE2; GB:BA000018; PID:gl3701517; PIDN:BAB42811.1; GSPDB:GN000001
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA1543

Query Match 83.8%; Score 31; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 98 HHPH 101

RESULT 24
S25846
homeotic protein Hox A4, testicular - mouse (fragment)
N;Alternate names: homeotic protein Hox 1.4
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 06-Feb-1998 #text_change 16-Aug-2004
C;Accession: S25847; S25846
R;Rubin, M.R.; Nguyen-Huu, M.C.
DNA Seq. 1, 329-334, 1991
A;Title: Murine embryonic spinal cord and adult testis Hox-1.4 cDNAs are identical 3' to 5' ends
A;Reference number: S25846; MUID:92190549; PMID:1686835
A;Accession: S25847
A;Molecule type: DNA
A;Residues: 1-28 <RUB1>
A;Cross-references: EMBL:X17346; NID:g51375; PID:g51376
A;Accession: S25846

A;Molecule type: mRNA
A;Residues: 27-155 <RUB2>
A;Cross-references: EMBL:X13538; NID:g51373; PIDN:CAA31889.1; PID:g51374
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;10-66/Domain: homeobox homology <HOX>

Query Match 83.8%; Score 31; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
|||
Db 98 HHPH 101

RESULT 25
T31664
DAP-kinase homolog COS1.3 - sea squirt (Ciona intestinalis) (fragment)
C;Species: Ciona intestinalis
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31664
R;Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z21050
A;Accession: T31664
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-156 <BIR>
A;Cross-references: UNIPROT:Q94426; EMBL:Z80904; NID:e1007749; PID:e274125; PIDN:CAB0258
C;Genetics:
A;Introns: 75/2; 131/3

Query Match 83.8%; Score 31; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
|||
Db 52 HHPH 55

RESULT 26
I58201
MHC class I antigen - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I58201
R;Lalanne, J.
Nucleic Acids Res. 10, 1039-1049, 1982
A;Title: Comparison of nucleotide sequences of mRNAs belonging to the mouse h-2 multigen
A;Reference number: I58201; MUID:82150234; PMID:6278432
A;Accession: I58201
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-164 <RES>
A;Cross-references: UNIPROT:Q31156; GB:J00395; NID:g199338; PIDN:AAA39579.1; PID:g387446
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;13-78/Domain: immunoglobulin homology <IMM>

Query Match 83.8%; Score 31; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
|||
Db 158 HHPH 161

RESULT 27
T13554
hypothetical protein 14 - Bacillus phage phi-105
C;Species: Bacillus phage phi-105

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13554
R;Kobayashi, K.; Okamura, K.; Inoue, T.; Sato, T.; Kobayashi, Y.
submitted to the EMBL Data Library, July 1998
A;Description: Complete nucleotide sequence of Bacillus subtilis phage phi-105.
A;Reference number: Z17688
A;Accession: T13554
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-171 <KOB>
A;Cross-references: UNIPROT:Q9ZXC1; EMBL:AB016282; PIDN:BAA36671.1
C;Superfamily: Bacillus phage phi-105 hypothetical protein 14

Query Match 83.8%; Score 31; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
|||
Db 23 HHPH 26

RESULT 28
T36013
probable integral membrane protein - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36013
R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21574
A;Accession: T36013
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-178 <SEE>
A;Cross-references: UNIPROT:Q8CK12; EMBL:AL096839; PIDN:CAB50766.1; GSPDB:GN00070; SCOEDB
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCC22.23c

Query Match 83.8%; Score 31; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4
|||
Db 150 HHPH 153

RESULT 29
B49773
ecdysone-dependent cuticle protein EDG-84 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C;Accession: B49773
R;Apple, R.T.; Fristrom, J.W.
Dev. Biol. 146, 569-582, 1991
A;Title: 20-Hydroxyecdysone is required for, and negatively regulates, transcription of I
A;Reference number: A49773; MUID:91323677; PMID:1713868
A;Accession: B49773
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-188 <APP>
A;Cross-references: UNIPROT:P27780; GB:M71249; NID:g157324; PIDN:AAA28501.1; PID:g157325
C;Genetics:
A;Gene: FlyBase:Edg84A
A;Cross-references: FlyBase:FBgn0000552

Query Match 83.8%; Score 31; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPH 4

Db 179 HHPH 182

||||

RESULT 30

E83864

thioredoxin BH1717 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: E83864

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: E83864

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-193 <STO>

A;Cross-references: UNIPROT:Q9KC57; GB:AP001512; GB:BA000004; NID:gl0174030; PIDN:BA054

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH1717

Query Match 83.8%; Score 31; DB 2; Length 193;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 5 HHPH 8

RESULT 31

S50686

hypothetical protein YER183c - yeast (Saccharomyces cerevisiae)

C;Species: Saccharomyces cerevisiae

C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004

C;Accession: S50686

R;Dietrich, F.S.

submitted to the EMBL Data Library, December 1994

A;Description: The sequence of S. cerevisiae cosmids 9163 and 9132.

A;Reference number: S50431

A;Accession: S50686

A;Molecule type: DNA

A;Residues: 1-211 <DIE>

A;Cross-references: UNIPROT:P40099; EMBL:U18922; NID:g603405; PID:g603424; GSPDB:GN00009

C;Genetics:

A;Gene: SGD:FAU1; MIPS:YER183c

A;Cross-references: SGD:S0000985

A;Map position: 5R

Query Match 83.8%; Score 31; DB 2; Length 211;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 93 HHPH 96

RESULT 32

G87457

guanylate kinase [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: G87457

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: G87457

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-213 <STO>

A;Cross-references: UNIPROT:Q9A7N9; GB:AE005673; NID:gl3423091; PIDN:AAK23659.1; GSPDB:GN

C;Genetics:

A;Gene: CC1681

C;Superfamily: guanylate kinase; guanylate kinase homology

Query Match 83.8%; Score 31; DB 2; Length 213;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 5 HHPH 8

RESULT 33

JX0244

pyroglutamyl-peptidase I (EC 3.4.19.3) - Bacillus amyloliquefaciens

N;Alternate names: 5-oxoprolyl-peptidase; pyroglutamyl aminopeptidase

C;Species: Bacillus amyloliquefaciens

C;Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 09-Jul-2004

C;Accession: JX0244

R;Yoshimoto, T.; Shimoda, T.; Kitazono, A.; Kabashima, T.; Ito, K.; Tsuru, D.

J. Biochem. 113, 67-73, 1993

A;Title: Pyroglutamyl peptidase gene from Bacillus amyloliquefaciens: Cloning, sequencing

A;Reference number: JX0244; MUID:93203177; PMID:8095933

A;Accession: JX0244

A;Molecule type: DNA

A;Residues: 1-215 <YOS>

A;Cross-references: UNIPROT:P46107; DDBJ:D11035; NID:g216315; PIDN:BAA01791.1; PID:g21631

C;Superfamily: pyrrolidone-carboxylate peptidase (N-terminal pyroglutamyl peptidase)

C;Keywords: hydrolase; omega peptidase

Query Match 83.8%; Score 31; DB 2; Length 215;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 158 HHPH 161

RESULT 34

T06455

Myb26 protein - garden pea

C;Species: Pisum sativum (garden pea)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T06455

R;Uimari, A.; Strommer, J.

Plant J. 12, 1273-1284, 1997

A;Title: Myb26: A MYB-like protein of pea flowers with affinity for promoters of phenylp

A;Reference number: Z15690; MUID:98112025; PMID:9450341

A;Accession: T06455

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-217 <UIM>

A;Cross-references: UNIPROT:P93474; EMBL:Y11105; NID:gl841474; PIDN:CAA71992.1; PID:gl841

C;Superfamily: Arabidopsis myb-related protein 2; myb DNA-binding repeat homology

C;Keywords: DNA binding

F;11-63/Domain: myb DNA-binding repeat homology <MYB>

F;64-113/Domain: myb DNA-binding repeat homology <MYB1>

Query Match 83.8%; Score 31; DB 2; Length 217;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 143 HHPH 146

RESULT 35
I59173
glutamate decarboxylase - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I59173
R;Bond, R.W.; Wyborski, R.J.; Gottlieb, D.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 8771-8775, 1990
A;Title: Developmentally regulated expression of an exon containing a stop codon in the
A;Reference number: I59173; MUID:91062362; PMID:2247446
A;Accession: I59173
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-223 <RES>
A;Cross-references: UNIPROT:Q63211; GB:M38350; NID:g204231; PIDN:AAA41185.1; PID:g204232
C;Superfamily: human glutamate decarboxylase

Query Match 83.8%; Score 31; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
Db 141 HHPH 144

RESULT 36
S35735
DNA gyrase chain A - Spiroplasma citri (fragment)
C;Species: Spiroplasma citri
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S35735
R;Ye, F.; Laigret, F.; Bove, J.
submitted to the EMBL Data Library, December 1992
A;Description: Nucleotide sequence and genetic organization at the replication origin (c
A;Reference number: S35732
A;Accession: S35735
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-227 <YEF>
A;Cross-references: UNIPROT:P34030; EMBL:Z19108; NID:g49345; PIDN:CAA79524.1; PID:g49349
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase
C;Keywords: ATP; DNA binding
F;7-227/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (frag

Query Match 83.8%; Score 31; DB 2; Length 227;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 84 YHPHG 88

RESULT 37
T40889
hypothetical protein SPCC1235.15 - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C;Accession: T40889
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z21954
A;Accession: T40889
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-228 <WOO>
A;Cross-references: EMBL:AL031764; PIDN:CAA21119.1; GSPDB:GN00068; SPDB:SPCC1235.15
A;Experimental source: strain 972h-; cosmid c1235
C;Genetics:
A;Gene: SPDB:SPCC1235.15

A;Map position: 3
C;Superfamily: Caenorhabditis elegans hypothetical protein K07B1.4

Query Match 83.8%; Score 31; DB 2; Length 228;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 118 YHPHG 122

RESULT 38
S65426
pyruvate decarboxylase (EC 4.1.1.1) - fava bean (fragment)
C;Species: Vicia faba (fava bean)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 12-Jul-2004
C;Accession: S65426
R;Muecke, U.; Wohlfarth, T.; Fiedler, U.; Baeumlein, H.; Ruecknagel, K.P.; Koenig, S.
Eur. J. Biochem. 237, 373-382, 1996
A;Title: Pyruvate decarboxylase from Pisum sativum. Properties, nucleotide and amino acid
A;Reference number: S65423; MUID:96215432; PMID:8647075
A;Accession: S65426
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-228 <MUE>
A;Cross-references: UNIPROT:Q7M228
C;Superfamily: thiamin pyrophosphate-binding domain homology
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 83.8%; Score 31; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
Db 95 HHPH 98

RESULT 39
H82360
conserved hypothetical protein VC0127 [imported] - Vibrio cholerae (strain N16961 serogro
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: H82360
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: H82360
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-232 <HBI>
A;Cross-references: UNIPROT:Q9KVL5; GB:AE004103; GB:AE003852; NID:g9654523; PIDN:AAF93304
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0127
A;Map position: 1

Query Match 83.8%; Score 31; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH .4
Db 27 HHPH 30

RESULT 40
AB0259
probable phage minor tail protein YP02123 [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB0259
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AB0259
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-236 <KUR>
A;Cross-references: UNIPROT:Q8ZEN8; GB:AL590842; PIDN:CAC90934.1; PID:gl5980130; GSPDB:G
C;Genetics:
A;Gene: YPO2123
C;Superfamily: phage lambda tail assembly protein K

Query Match 83.8%; Score 31; DB 2; Length 236;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||:|
Db 205 HHPYG 209

Search completed: June 15, 2005, 14:22:47
Job time : 21.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-8
Perfect score: 37
Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_eprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	37	100.0	72	2	Q8CLE3	Q8cle3 yersinia pe
2	37	100.0	189	2	Q6E2P7	Q6e2p7 porana comm
3	37	100.0	192	2	Q6E2S2	Q6e2s2 ipomoea pes
4	37	100.0	198	2	Q9W4B3	Q9w4b3 drosophila
5	37	100.0	241	2	Q8TDG7	Q8tdg7 homo sapien
6	37	100.0	246	2	Q87QR5	Q87qr5 vibrio para
7	37	100.0	273	1	CXB5_HUMAN	Q95377 homo sapien
8	37	100.0	284	2	Q9VAN2	Q9van2 drosophila
9	37	100.0	287	2	Q9W7B8	Q9w7b8 brachydanio
10	37	100.0	294	2	Q65F52	Q65f52 bacillus li
11	37	100.0	295	2	Q7QID0	Q7qid0 anopheles g
12	37	100.0	312	2	Q68DR3	Q68dr3 homo sapien
13	37	100.0	331	2	Q67PX1	Q67px1 symbiobacte
14	37	100.0	340	2	Q8DJ31	Q8dj31 synechococc
15	37	100.0	348	2	Q9X6C8	Q9x6c8 thermus bro
16	37	100.0	349	2	Q8GEA7	Q8gea7 thermus sp.
17	37	100.0	349	2	Q9X6D3	Q9x6d3 thermus the
18	37	100.0	349	2	Q746I4	Q746i4 thermus the
19	37	100.0	356	2	O73679	O73679 brachydanio
20	37	100.0	356	2	Q98UK5	Q98uk5 brachydanio
21	37	100.0	396	1	HRG_BOVIN	P33433 bos taurus
22	37	100.0	419	1	GSC_DROME	P54366 drosophila
23	37	100.0	428	2	Q9PW_C8	Q9pwc8 brachydanio
24	37	100.0	430	2	Q98SS0	Q98ss0 brachydanio
25	37	100.0	430	2	Q98SS3	Q98ss3 brachydanio
26	37	100.0	431	2	Q98SS2	Q98ss2 brachydanio
27	37	100.0	432	2	Q98SS1	Q98ss1 brachydanio
28	37	100.0	469	2	Q8MQS2	Q8mqs2 drosophila
29	37	100.0	491	1	SPEA_BACHD	Q9k9k5 bacillus ha
30	37	100.0	509	2	Q9VIW0	Q9viw0 drosophila
31	37	100.0	510	2	Q9ESB2	Q9esb2 rattus norv

32	37	100.0	515	2	Q99PS7	Q99ps7 rattus norv
33	37	100.0	524	2	Q80XK4	Q80xk4 mus musculu
34	37	100.0	525	1	HRG_HUMAN	P04196 homo sapien
35	37	100.0	525	2	Q99PS5	Q99ps5 mus musculu
36	37	100.0	525	2	Q99PS6	Q99ps6 mus musculu
37	37	100.0	525	2	Q99PS8	Q99ps8 rattus norv
38	37	100.0	525	2	Q9ESB3	Q9esb3 mus musculu
39	37	100.0	526	1	HRG_RABIT	Q28640 oryctolagus
40	37	100.0	536	2	Q6YK32	Q6yk32 mus musculu
41	37	100.0	536	2	Q6YKA2	Q6yka2 mus musculu
42	37	100.0	539	2	Q9NQA4	Q9nqa4 homo sapien
43	37	100.0	539	2	Q7TNS7	Q7tns7 mus musculu
44	37	100.0	539	2	Q9JHS6	Q9jhs6 rattus norv
45	37	100.0	539	2	Q9QYV9	Q9qyv9 rattus norv

ALIGNMENTS

RESULT 1

Q8CLE3	ID	Q8CLE3	PRELIMINARY;	PRT;	72 AA.
AC	Q8CLE3;				
DT	01-MAR-2003	(TReMBLrel. 23, Created)			
DT	01-MAR-2003	(TReMBLrel. 23, Last sequence update)			
DT	01-MAR-2003	(TReMBLrel. 23, Last annotation update)			
DE	Hypothetical.				
GN	OrderedLocusNames=y1535;				
OS	Yersinia pestis.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Yersinia.				
OX	NCBI_TaxID=632;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=KIMS / Biovar Mediaevalis;				
RX	MEDLINE=22137863; PubMed=12142430;				
RX	DOI=10.1128/JB.184.16.4601-4611.2002;				
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,				
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,				
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,				
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,				
RA	Perry R.D.;				
RT	"Genome sequence of Yersinia pestis KIM."				
RL	J. Bacteriol. 184:4601-4611(2002).				
DR	EMBL; AE013756; AAM85107.1; --				
KW	Hypothetical protein.				
SQ	SEQUENCE 72 AA; 7848 MW; D2C5B430EE62B0DA CRC64;				

Query Match 100.0%; Score 37; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	HHPHG 5
Db	42	HHPHG 46

RESULT 2

Q6E2P7	ID	Q6E2P7	PRELIMINARY;	PRT;	189 AA.
AC	Q6E2P7;				
DT	25-OCT-2004	(TReMBLrel. 28, Created)			
DT	25-OCT-2004	(TReMBLrel. 28, Last sequence update)			
DT	25-OCT-2004	(TReMBLrel. 28, Last annotation update)			
DE	Ribosomal protein L2 (Fragment).				
GN	Name=rpL2;				
OS	Porana commixta.				
OG	Chloroplast.				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;				
OC	lamids; Solanales; Convolvulaceae; Dichondreae; Porana.				
OX	NCBI_TaxID=197434;				
RN	[1]				

RP SEQUENCE FROM N.A.
RA Stefanovic S., Olmstead R.G.;
RT "Testing the Phylogenetic Position of a Parasitic Plant (Cuscuta,
RT Convolvulaceae, Asteridae): Bayesian Inference and the Parametric
RT Bootstrap on Data Drawn from Three Genomes.";
RL Syst. Biol. 53:384-399(2004).
DR EMBL; AY596770; AAT69100.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0015934; C:large ribosomal subunit; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR002171; Ribosomal_L2.
DR InterPro; IPR005880; Ribosomal_L2_b/o.
DR InterPro; IPR008991; Transl_SH3_like.
DR Pfam; PF00181; Ribosomal_L2; 1.
DR Pfam; PF03947; Ribosomal_L2_C; 1.
DR TIGRFAMS; TIGR01171; rplB_bact; 1.
KW Chloroplast; Ribosomal protein.
FT NON_TER 1 1
FT NON_TER 189 189
SQ SEQUENCE 189 AA; 20451 MW; 2C81496B6779C7B5 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 181 HHPHG 185

RESULT 3
Q6E2S2
ID Q6E2S2 PRELIMINARY; PRT; 192 AA.
AC Q6E2S2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribosomal protein L2 (Fragment).
GN Name=rpl2;
OS Ipomoea pes-tigridis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Convolvulaceae; Ipomoeae; Ipomoea.
OX NCBI_TaxID=89657;
RN [1]
RP SEQUENCE FROM N.A.
RA Stefanovic S., Olmstead R.G.;
RT "Testing the Phylogenetic Position of a Parasitic Plant (Cuscuta,
RT Convolvulaceae, Asteridae): Bayesian Inference and the Parametric
RT Bootstrap on Data Drawn from Three Genomes.";
RL Syst. Biol. 53:384-399(2004).
DR EMBL; AY596745; AAT69075.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0015934; C:large ribosomal subunit; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR002171; Ribosomal_L2.
DR InterPro; IPR005880; Ribosomal_L2_b/o.
DR InterPro; IPR008991; Transl_SH3_like.
DR Pfam; PF00181; Ribosomal_L2; 1.
DR Pfam; PF03947; Ribosomal_L2_C; 1.
DR TIGRFAMS; TIGR01171; rplB_bact; 1.
KW Chloroplast; Ribosomal protein.

FT NON_TER 1 1
FT NON_TER 192 192
SQ SEQUENCE 192 AA; 20573 MW; 69D683AEF55FEE0D CRC64;

Query Match 100.0%; Score 37; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 179 HHPHG 183

RESULT 4
Q9W4B3
ID Q9W4B3 PRELIMINARY; PRT; 198 AA.
AC Q9W4B3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG15781-PA.
GN ORFNames=CG15781;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
"Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
SEQUENCE FROM N.A.
MEDLINE=22426070; PubMed=12537573;
Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
Ashburner M., Celniker S.E.;
"The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
SEQUENCE FROM N.A.
MEDLINE=22426069; PubMed=12537572;
Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
Lewis S.E.;
"Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
SEQUENCE FROM N.A.
FlyBase;
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A.
FlyBase;
Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
EMBL; AE003434; AAF46043.1; -.
IntAct; Q9W4B3; -.
FlyBase; FBgn0029774; CGI5781.
SEQUENCE 198 AA; 21401 MW; A0A6CAD90F6AA252 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 153 HHPHG 157
RESULT 5
Q8TDG7 PRELIMINARY; PRT; 241 AA.
AC Q8TDG7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CLL-associated antigen KW-6 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Krackhardt A.M., Witzens M., Harig S., Hodi F.S., Zauls A.J.,
Chessia M., Barrett P., Gribben J.G.;
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF432219; AAL99925.1; -.
FT NON_TER 1
SQ SEQUENCE 241 AA; 25884 MW; 33C30E6088012685 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 241;

Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 232 HHPHG 236
RESULT 6
Q87QR5 PRELIMINARY; PRT; 246 AA.
AC Q87QR5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VP1084.
GN OrderedLocusNames=VP1084;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
"Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005076; BAC59347.1; -.
KW Complete proteome.
SQ SEQUENCE 246 AA; 27594 MW; 3F130DD284361E45 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 14 HHPHG 18
RESULT 7
CXBS_HUMAN STANDARD; PRT; 273 AA.
ID CXBS_HUMAN
AC Q95377; Q9UPA3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Gap junction beta-5 protein (Connexin 31.1) (Cx31.1).
GN Name=GJB5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99057343; PubMed=9843209; DOI=10.1038/3840;
RA Richard G., Smith L.E., Bailey R.A., Itin P., Hohl D.,
RA Epstein E.H. Jr., DiGiovanna J.J., Compton J.G., Bale S.J.;
"Mutations in the human connexin gene GJB3 cause erythrodermia
variabilis.";
RL Nat. Genet. 20:366-369(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Xia J.-H., Liu C.-Y., Zheng D., Pan Q., Xie W.;
"Molecular cloning of human connexin 31 and 31.1";
Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Pearce A.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=124777932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carminci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low MW diffuse from one cell to a neighboring cell.
CC -!- SUBUNIT: A connexon is composed of a hexamer of connexins.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the connexin family. Beta-type (group I)
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF099731; AAC95472.1; --
CC EMBL; AF052693; AAD18005.1; --
CC EMBL; AL121988; CAB90271.1; --
CC EMBL; BC004379; AAH04379.1; --
CC Genew; HGNC:4287; GJB5.
CC MIM; 604493; --
CC GO; GO:0008544; P:epidermal differentiation; TAS.
CC InterPro; IPR000500; Connexin.
CC InterPro; IPR002270; Connexin311.
CC Pfam; PF00029; Connexin; 1.
CC PRINTS; PR00206; CONNEXIN.
CC SMART; SM00037; CNX; 1.
CC PROSITE; PS00407; CONNEXINS_1; 1.
CC PROSITE; PS00408; CONNEXINS_2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 20 Cytoplasmic (Potential).
FT TRANSMEM 21 40 Potential.
FT DOMAIN 41 75 Extracellular (Potential).
FT TRANSMEM 76 98 Potential.
FT DOMAIN 99 126 Cytoplasmic (Potential).
FT TRANSMEM 127 149 Potential.
FT DOMAIN 150 187 Extracellular (Potential).
FT TRANSMEM 188 210 Potential.
FT DOMAIN 211 273 Cytoplasmic (Potential).
SQ SEQUENCE 273 AA; 31088 MW; C1B9F1A4A2E070BB CRC64;
Query Match 100.0%; Score 37; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HHPHG 5
|||||

Db 229 HHPHG 233
RESULT 8
Q9VAN2 PRELIMINARY; PRT; 284 AA.
AC Q9VAN2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CG14509-PA.
GN ORFNames=CG14509;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.

RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.

RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.

RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003768; AAF56872.2; -.
DR FlyBase; FBgn0039647; CG14509.
SQ SEQUENCE 284 AA; 31482 MW; A94AA7075D40446C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 240 HHPHG 244

RESULT 9
Q9W7B8 PRELIMINARY; PRT; 287 AA.
AC Q9W7B8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcription factor Tcf3b (Fragment).
GN Name=tcf7l1b; Synonyms=tcf3b;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=99376682; PubMed=10446273; DOI=10.1016/S0925-4773(99)00101-X;
RA Dorsky R.I., Snyder A., Cretekos C.J., Grunwald D.J., Geisler R.,
RA Haffter P., Moon R.T., Raible D.W.;
RT "Maternal and embryonic expression of zebrafish left1.";
RL Mech. Dev. 86:147-150(1999).
DR HSSP; P27782; 2LEF.
DR ZFIN; ZDB-GENE-991110-10; tcf7l1b.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009071; HMG-box.
DR InterPro; IPR00910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS50118; HMG_BOX_2; 1.
FT NON_TER 1
SQ SEQUENCE 287 AA; 31586 MW; 8573199746D91F0F CRC64;

Query Match 100.0%; Score 37; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 12 HHPHG 16

RESULT 10
Q65F52 PRELIMINARY; PRT; 294 AA.
AC Q65F52;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE YuxN (Homeodomain-like).
GN Name=yuxN; ORFNames=BL00748, BLi03485;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AE017333; AAU42312.1; -.
DR EMBL; CP000002; AAU24943.1; -.
KW Nuclear protein.
SQ SEQUENCE 294 AA; 34659 MW; 8C74565A1ABA478C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 290 HHPHG 294

RESULT 11
Q7QID0 PRELIMINARY; PRT; 295 AA.
AC Q7QID0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AgCP3965 (Fragment).
GN Name=agCG52960; ORFNames=ENSANGG00000017877;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;

RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBDJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBDJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008807; EAA04166.1; -.
DR HSSP; Q99958; 1D5V.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001766; TF_Fork_head.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00039; FORK_HEAD_3; 1.
FT NON_TER 295
SQ SEQUENCE 295 AA; 32310 MW; CBB8AF12C2DA2970 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 272 HHPHG 276

RESULT 12
Q68DR3 PRELIMINARY; PRT; 312 AA.
ID Q68DR3
AC Q68DR3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp779H1622 (Fragment).
GN Name=DKFZp779H1622;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBDJ databases.
DR EMBL; CR749302; CAH18157.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 312 AA; 35124 MW; 08522797450AFA1F CRC64;

Query Match 100.0%; Score 37; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 171 HHPHG 175

RESULT 13
Q67PX1 PRELIMINARY; PRT; 331 AA.
ID Q67PX1
AC Q67PX1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=STH1287;

OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium
RT thermophilum.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AP006840; BAD40272.1; -.
KW Hypothetical protein.
SQ SEQUENCE 331 AA; 36263 MW; 2EF2D7BD61A84880 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 149 HHPHG 153

RESULT 14
Q8DJ31 PRELIMINARY; PRT; 340 AA.
ID Q8DJ31
AC Q8DJ31;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tll1398 protein.
GN OrderedLocusNames=tll1398;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005373; BAC08950.1; -.
DR InterPro; IPR002762; CbiX.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01903; CbiX; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 340 AA; 38706 MW; A9363ED422C0BADB CRC64;

Query Match 100.0%; Score 37; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 317 HHPHG 321

RESULT 15
Q9X6C8 PRELIMINARY; PRT; 348 AA.
ID Q9X6C8
AC Q9X6C8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase homolog.
GN Name=galt;

OS Thermus brockianus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=56956;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ITI360;
RX MEDLINE=99402735; PubMed=10473401;
RA Fridjonsson O., Watzlawick H., Gehweiler A., Rohrhirsch T., Mattes R.;
RT "Cloning of the gene encoding a novel thermostable alpha-galactosidase
from Thermus brockianus ITI360.";
RL Appl. Environ. Microbiol. 65:3955-3963(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ITI360;
RX MEDLINE=20203878; PubMed=10741834;
RA Fridjonsson O., Watzlawick H., Mattes R.;
RT "The structure of the alpha-galactosidase gene loci in Thermus
brockianus ITI360 and Thermus thermophilus TH125.";
RL Extremophiles 4:23-33(2000).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
CC alpha-D-glucose 1-phosphate + UDP-galactose.
CC -!- PATHWAY: Galactose metabolism; second step.
CC -!- SIMILARITY: Belongs to the galactose-1-phosphate
CC uridylyltransferase family 1.
DR EMBL; AF135398; AAD33669.1; -.
DR HSSP; P09148; 1GUQ.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . .; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . .; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; Galp_UDPtransf1.
DR InterPro; IPR011573; Galp_Utransf.
DR ProDom; PD005051; Galp_Utransf; 2.
DR TIGRFAMs; TIGR00209; galt_1; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF_I; 1.
KW Carbohydrate metabolism; Galactose metabolism; Nucleotidyltransferase;
KW Transference.
SQ SEQUENCE 348 AA; 39855 MW; BF6CA7EEA85DC904 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 183 HHPHG 187

RESULT 16
Q8GEA7
ID Q8GEA7 PRELIMINARY; PRT; 349 AA.
AC Q8GEA7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase-like protein.
OS Thermus sp. IB-21.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=206164;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IB-21;
RA Kang S.K., Cho K.K., Ahn J.K., Choi Y.J.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
CC alpha-D-glucose 1-phosphate + UDP-galactose.

CC -!- PATHWAY: Galactose metabolism; second step.
CC -!- SIMILARITY: Belongs to the galactose-1-phosphate
CC uridylyltransferase family 1.
DR EMBL; AY130259; AAN05445.1; -.
DR HSSP; P09148; 1GUQ.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . .; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . .; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; Galp_UDPtransf1.
DR InterPro; IPR011573; Galp_Utransf.
DR InterPro; IPR005849; Galp_Utransf_N.
DR Pfam; PF01087; Galp_UDP_transf; 1.
DR PIRSF; PIRSF000808; Galt; 1.
DR ProDom; PD005051; Galp_Utransf; 2.
DR TIGRFAMs; TIGR00209; galt_1; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF_I; 1.
KW Carbohydrate metabolism; Galactose metabolism; Nucleotidyltransferase;
KW Transference.
SQ SEQUENCE 349 AA; 39736 MW; DF8A239FA7C59607 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 183 HHPHG 187
RESULT 17
Q9X6D3
ID Q9X6D3 PRELIMINARY; PRT; 349 AA.
AC Q9X6D3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase.
GN Name=galt;
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH125;
RX MEDLINE=20203878; PubMed=10741834;
RA Fridjonsson O., Watzlawick H., Mattes R.;
RT "The structure of the alpha-galactosidase gene loci in Thermus
brockianus ITI360 and Thermus thermophilus TH125.";
RL Extremophiles 4:23-33(2000).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
CC alpha-D-glucose 1-phosphate + UDP-galactose.
CC -!- PATHWAY: Galactose metabolism; second step.
CC -!- SIMILARITY: Belongs to the galactose-1-phosphate
CC uridylyltransferase family 1.
DR EMBL; AF135399; AAD32629.1; -.
DR HSSP; P09148; 1GUQ.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . .; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . .; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; Galp_UDPtransf1.
DR InterPro; IPR011573; Galp_Utransf.
DR InterPro; IPR005850; Galp_Utransf_C.
DR InterPro; IPR005849; Galp_Utransf_N.
DR Pfam; PF01087; Galp_UDP_transf; 1.
DR Pfam; PF02744; Galp_UDP_tr_C; 1.
DR PIRSF; PIRSF000808; Galt; 1.
DR ProDom; PD005051; Galp_Utransf; 2.
DR TIGRFAMs; TIGR00209; galt_1; 1.
KW Carbohydrate metabolism; Galactose metabolism; Nucleotidyltransferase;
KW Transference.

SQ SEQUENCE 349 AA; 39451 MW; 0B2284608704339A CRC64;

Query Match 100.0%; Score 37; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||
Db 183 HHPHG 187

RESULT 18
Q746I4 PRELIMINARY; PRT; 349 AA.

AC Q746I4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10).
GN OrderedLocusNames=JTP0071;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OG Plasmid pT27.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064768;

RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartsch T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
thermophilus.";
RL Nat. Biotechnol. 22:547-553 (2004).
CC -|- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
CC alpha-D-glucose 1-phosphate + UDP-galactose.
CC -|- PATHWAY: Galactose metabolism; second step.
CC -|- SIMILARITY: Belongs to the galactose-1-phosphate
CC uridylyltransferase family 1.
DR EMBL; AE017222; AAS82401.1; -.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . .; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . .; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; GalP_UDPtransf1.
DR InterPro; IPR011573; GalP_Utransf.
DR InterPro; IPR005850; GalP_Utransf_C.
DR InterPro; IPR005849; GalP_Utransf_N.
DR Pfam; PF02744; GalP_UDP_trf_C; 1.
DR Pfam; PF02744; GalP_UDP_trf_C; 1.
DR PIRSF; PIRSF00808; GalT; 1.
DR ProDom; PD005051; GalP_Utransf; 2.
DR TIGRFAMs; TIGR00209; galT_1; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF_I; 1.
KW Carbohydrate metabolism; Complete proteome; Galactose metabolism;
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 349 AA; 39576 MW; 2045DC5BE440FA76 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||
Db 183 HHPHG 187

RESULT 19
O73679 PRELIMINARY; PRT; 356 AA.

AC O73679;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcription factor Val.
GN Name=mafB; Synonyms=valentino;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98165393; PubMed=9425134;
RA Moens C.B., Cordes S.P., Giorgianni M.W., Barsh G.S., Kimmel C.B.;
RT "Equivalence in the genetic control of hindbrain segmentation in fish
RT and mouse.";
RL Development 125:381-391 (1998).
CC -|- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AF006641; AAC18821.1; -.
DR HSSP; O54790; 1K1V.
DR ZFIN; ZDB-GENE-980526-515; mafB.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk_transcr_DNA.
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR004826; TF_Maf.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 356 AA; 40243 MW; 07420DB0F6CD08F1 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||
Db 219 HHPHG 223

RESULT 20
Q98UK5 PRELIMINARY; PRT; 356 AA.

AC Q98UK5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcription factor MafB.
GN Name=mafB; Synonyms=mafB;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21064923; PubMed=11134968;
RA Kajihara M., Kawauchi S., Kobayashi M., Ogino H., Takahashi S.,
RA Yasuda K.;
RT "Isolation, characterization, and expression analysis of zebrafish
RT large Mafs.";
RL J. Biochem. 129:139-146 (2001).
CC -|- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AB006322; BAB21102.1; -.
DR HSSP; O54790; 1K1V.
DR ZFIN; ZDB-GENE-980526-515; mafB.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk_transcr_DNA.
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR004826; TF_Maf.
DR Pfam; PF03131; bZIP_Maf; 1.


```
DR SMART; SM00338; BRL2; 1.
DR PROSITE; PS50217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 356 AA; 40233 MW; DE4C96B62C058865 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 219 HHPHG 223

RESULT 21
HRG_BOVIN
ID HRG_BOVIN STANDARD; PRT; 396 AA.
AC P33433;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein (Histidine-proline rich glycoprotein)
DE (HPRG) (Fragments).
GN Name=HRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=93351678; PubMed=8348977; DOI=10.1016/0014-5793(93)80945-Q;
RA Soerensen C.B., Krogh-Pedersen H., Petersen T.E.;
RT "Determination of the disulphide bridge arrangement of bovine
RT histidine-rich glycoprotein.";
RL FEBS Lett. 328:285-290(1993).
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 12 tandem repetitions of a 5-
CC residue sequence (GHHPH, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.
DR InterPro; IPR000010; Prot_inh_cystat.
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat.
FT DOMAIN 1 102 Cystatin-like 1.
FT DOMAIN 103 169 Cystatin-like 2.
FT DOMAIN 191 238 Pro-rich.
FT DOMAIN 243 368 His/Pro-rich.
FT DISULFID 7 375
FT DISULFID 56 67
FT DISULFID 77 92
FT DISULFID 123 297
FT DISULFID 137 160
FT DISULFID 212 242
FT NON_CONS 52 53
FT CARBOHYD 70 70 N-linked (GlcNAc. . .).
FT NON_CONS 71 72
FT NON_CONS 78 79
FT CARBOHYD 91 91 N-linked (GlcNAc. . .).
FT NON_CONS 103 104
FT CARBOHYD 122 122 N-linked (GlcNAc. . .).
FT NON_CONS 163 164
FT CARBOHYD 220 220 N-linked (GlcNAc. . .).
FT NON_CONS 263 264
FT NON_CONS 303 304
FT VARIANT 86 86 S -> R.
```

```
FT VARIANT 309 322 S -> Q.
FT VARIANT 322 322 H -> Y.
SQ SEQUENCE 396 AA; 44470 MW; 128A8223499DE6FC CRC64;

Query Match 100.0%; Score 37; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 264 HHPHG 268

RESULT 22
GSC_DROME
ID GSC_DROME STANDARD; PRT; 419 AA.
AC P54366; Q9VPR9;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Homeobox protein goosecoid.
GN Name=Gsc; ORFNames=CG2851;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=96202483; PubMed=8625850;
RA Goriely A., Stella M., Coffinier C., Kessler D., Mailhos C.,
RA Dessain S., Desplan C.;
RT "A functional homologue of goosecoid in Drosophila.";
RL Development 122:1641-1650(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96272167; PubMed=8670808;
RA Hahn M., Jackle H.;
RT "Drosophila goosecoid participates in neural development but not in
RT body axis formation.";
RL EMBO J. 15:3077-3084(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: Appears to regulate regional development of specific
CC tissues. Can rescue axis polarity in UV-radiated Xenopus embryos.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: In early embryo development, expression
CC confined to two regions; a horseshoe-like pattern across the
CC dorsal side which is destined to form the brain hemispheres and a
CC second domain which invaginates inside the stomodeum and which, is
CC fated to form the foregut, ring gland and stomatogastric nervous
CC system (SNS).
CC -!- SIMILARITY: Belongs to the paired homeobox family. Bicoid
CC subfamily.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X95420; CAA64699.1; -.
DR EMBL; U52968; AAB17948.1; -.
DR EMBL; AE003589; AAF51473.1; -.
DR PIR; S70617; S70617.
DR HSSP; P06601; 1FJL.
DR IntAct; P54366; -.
DR TRNSFAC; T04041; -.
DR FlyBase; FBgn0010323; Gsc.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR007104; Paired_homeo.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Developmental protein; DNA-binding; Homeobox; Nuclear protein.
FT DOMAIN 104 107 Poly-Ala.
FT DOMAIN 164 169 Poly-Ser.
FT DOMAIN 195 199 Poly-Ala.
FT DNA BIND 286 345 Homeobox.
SQ SEQUENCE 419 AA; 44949 MW; 851A4C46AA861FB9 CRC64;

Query Match 100.0%; Score 37; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 253 HHPHG 257
RESULT 23
Q9PWC8
ID Q9PWC8 PRELIMINARY; PRT; 428 AA.
AC Q9PWC8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Egg envelope protein ZP2.
GN Name=ZP2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99326148; PubMed=10395930; DOI=10.1016/S0167-4781(99)00066-4;
RA Wang H., Gong Z.;
RT "Characterization of two zebrafish cDNA clones encoding egg envelope
RT proteins ZP2 and ZP3.";
RL Biochim. Biophys. Acta 1446:156-160(1999).
DR EMBL; AF095456; AAD49112.1; -.
DR ZFIN; ZDB-GENE-991129-5; ZP2.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; Trefoil; 1.
DR Pfam; PF00100; Zona pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00018; PD; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Envelope protein.
SQ SEQUENCE 428 AA; 46804 MW; 46E8825ACA17C6A7 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 37 HHPHG 41
RESULT 24
Q98SS0
ID Q98SS0 PRELIMINARY; PRT; 430 AA.
AC Q98SS0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Egg envelope protein ZP2.
GN Name=ZP2.2; Synonyms=ZP2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21017556; PubMed=11144219;
RX DOI=10.1002/1098-2795(200101)58:1<4::AID-MRD2>3.0.CO;2-P;
RA Mold D.E., Kim I.F., Tsai C.-M., Lee D., Chang C.-Y., Huang R.C.;
RT "Cluster of genes encoding the major egg envelope protein of
RT zebrafish.";
RL Mol. Reprod. Dev. 58:4-14(2001).
DR EMBL; AF331968; AAK16580.1; -.

DR	ZFIN; ZDB-GENE-010622-1; zp2.2.	DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR	GO; GO:0019031; C:viral envelope; IEA.	DE	Egg envelope protein ZP2 variant B.
DR	InterPro; IPR001507; Endoglin/CD105.	GN	Name=ZP2.3; Synonyms=ZP2;
DR	InterPro; IPR000519; P_trefoil.	OS	Brachydanio rerio (Zebrafish) (Danio rerio).
DR	Pfam; PF00088; Trefoil; 1.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DR	Pfam; PF00100; Zona_pellucida; 1.	OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
DR	PRINTS; PR00023; ZPELLUCIDA.	OC	Cyprinidae; Danio.
DR	SMART; SM00018; PD; 1.	OX	NCBI_TaxID=7955;
DR	SMART; SM00241; ZP; 1.	RN	[1]
DR	PROSITE; PS00682; ZP_DOMAIN; 1.	RP	SEQUENCE FROM N.A.
KW	Envelope protein.	RX	MEDLINE=21017556; PubMed=11144219;
SQ	SEQUENCE 430 AA; 47124 MW; EC54C4EC03A10B63 CRC64;	RX	DOI=10.1002/1098-2795(200101)58:1<4::AID-MRD2>3.0.CO;2-P;
Query Match 100.0%; Score 37; DB 2; Length 430;		RA	Mold D.E., Kim I.F., Tsai C.-M., Lee D., Chang C.-Y., Huang R.C.;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;		RT	"Cluster of genes encoding the major egg envelope protein of
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		RT	zebrafish.";
QY	1 HHPHG 5	RL	Mol. Reprod. Dev. 58:4-14(2001).
Db		DR	EMBL; AF331966; AAK16577.1; -.
	40 HHPHG 44	DR	ZFIN; ZDB-GENE-010622-2; zp2.3.
RESULT 25		DR	GO; GO:0019031; C:viral envelope; IEA.
Q98SS3		DR	InterPro; IPR001507; Endoglin/CD105.
ID	Q98SS3 PRELIMINARY; PRT; 430 AA.	DR	InterPro; IPR000519; P_trefoil.
AC	Q98SS3;	DR	Pfam; PF00088; Trefoil; 1.
DT	01-JUN-2001 (TrEMBLrel. 17, Created)	DR	PRINTS; PR00023; ZPELLUCIDA.
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	DR	SMART; SM00018; PD; 1.
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	DR	SMART; SM00241; ZP; 1.
DE	Egg envelope protein ZP2 variant A.	DR	PROSITE; PS00682; ZP_DOMAIN; 1.
GN	Name=ZP2.2; Synonyms=ZP2;	KW	Envelope protein.
OS	Brachydanio rerio (Zebrafish) (Danio rerio).	SQ	SEQUENCE 430 AA; 47124 MW; EC54C4EC03A10B63 CRC64;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Query Match 100.0%; Score 37; DB 2; Length 430;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	Best Local Similarity 100.0%; Pred. No. 1.8e+02;	
OC	Cyprinidae; Danio.	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21017556; PubMed=11144219;		
RX	DOI=10.1002/1098-2795(200101)58:1<4::AID-MRD2>3.0.CO;2-P;		
RA	Mold D.E., Kim I.F., Tsai C.-M., Lee D., Chang C.-Y., Huang R.C.;		
RT	"Cluster of genes encoding the major egg envelope protein of		
RT	zebrafish.";		
RL	Mol. Reprod. Dev. 58:4-14(2001).		
DR	EMBL; AF331965; AAK16578.1; -.		
DR	ZFIN; ZDB-GENE-010622-1; zp2.2.		
DR	GO; GO:0019031; C:viral envelope; IEA.		
DR	InterPro; IPR001507; Endoglin/CD105.		
DR	InterPro; IPR000519; P_trefoil.		
DR	Pfam; PF00088; Trefoil; 1.		
DR	Pfam; PF00100; Zona_pellucida; 1.		
DR	PRINTS; PR00023; ZPELLUCIDA.		
DR	SMART; SM00018; PD; 1.		
DR	SMART; SM00241; ZP; 1.		
DR	PROSITE; PS00682; ZP_DOMAIN; 1.		
KW	Envelope protein.		
SQ	SEQUENCE 430 AA; 47083 MW; 7544910DAC97806E CRC64;		
Query Match 100.0%; Score 37; DB 2; Length 430;			
Best Local Similarity 100.0%; Pred. No. 1.8e+02;			
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 HHPHG 5		
Db			
	40 HHPHG 44		
RESULT 26			
Q98SS2			
ID	Q98SS2 PRELIMINARY; PRT; 431 AA.		
AC	Q98SS2;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		

SQ SEQUENCE 432 AA; 47343 MW; B985118D36ED5B8C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||
Db 40 HHPHG 44

RESULT 28
Q8MQS2
ID Q8MQS2 PRELIMINARY; PRT; 469 AA.
AC Q8MQS2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GH11945p.
GN ORFNames=CG14509;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dreesnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY128418; AAM75011.1; -.
DR FlyBase; FBgn0039647; CG14509.
SQ SEQUENCE 469 AA; 51845 MW; BECEID5C4D18CC06 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||
Db 425 HHPHG 429

RESULT 29
SPEA_BACHD
ID _SPEA_BACHD STANDARD; PRT; 491 AA.
AC Q9K9K5;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Arginine decarboxylase (EC 4.1.1.19).
GN Name=speA; OrderedLocusNames=BH2640;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- FUNCTION: Catalyzes the formation of agmatine from arginine (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: L-arginine = agmatine + CO(2).

CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Putrescine biosynthesis from arginine; first step.
CC -!- PATHWAY: Spermidine biosynthesis from arginine; first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the Orn/Lys/Arg decarboxylase class-I
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001516; BAB06359.1; -.
DR PIR; H83979; H83979.
DR InterPro; IPR000310; Decarboxylasel.
DR InterPro; IPR008286; Decarboxylase_C.
DR InterPro; IPR011185; Lys decarb.
DR Pfam; PF01276; OKR_DC_1; 1.
DR Pfam; PF03711; OKR_DC_1_C; 1.
DR PIRSF; PIRSF005938; Lys decarb; 1.
DR PROSITE; PS00703; OKR_DC_1; 1.
KW Complete proteome; Decarboxylase; Lyase; Polyamine biosynthesis;
KW Putrescine biosynthesis; Pyridoxal phosphate; Spermidine biosynthesis.
FT BINDING 227 227 Pyridoxal phosphate (By similarity).
SQ SEQUENCE 491 AA; 53449 MW; 3C2ACC55C3C57F5D CRC64;

Query Match 100.0%; Score 37; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||||
Db 62 HHPHG 66

RESULT 30
Q9VIW0
ID Q9VIW0 PRELIMINARY; PRT; 509 AA.
AC Q9VIW0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CG10034-PA (Traffic jam).
GN ORFNames=CG10034;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradercky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=22954636; PubMed=14578908; DOI=10.1038/ncb1058;
RA Li M.A., Alls J.D., Avancini R.M., Koo K., Godt D.;
RT "The large Maf factor Traffic Jam controls gonad morphogenesis in
RT Drosophila.";
RL Nat. Cell Biol. 5:994-1000(2003).
DR EMBL; AE003663; AAF53804.2; -.
DR EMBL; AY325814; AAP88969.1; -.
DR HSSP; O54790; 1K1V.
DR IntAct; Q9VIW0; -.

DR FlyBase; FBgn0000964; tj.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk transcr_DNA.
DR InterPro; IPR004827; TF_5ZIP.
DR InterPro; IPR004826; TF_Maf.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
SQ SEQUENCE 509 AA; 54043 MW; 0BD429AFB0BCCF38 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 267 HHPHG 271

RESULT 31
Q9ESB2
ID Q9ESB2 PRELIMINARY; PRT; 510 AA.
AC Q9ESB2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name:Hrg;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lewis;
RX MEDLINE=20307726; PubMed=10849117;
RA Hulett M.D., Parish C.R.;
RT "Murine histidine-rich glycoprotein: cloning, characterization and
RT cellular origin.";
RL Immunol. Cell Biol. 78:280-287(2000).
DR EMBL; AF194029; AAG28417.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 510 AA; 57581 MW; 508E6E06AA2ED58E CRC64;

Query Match 100.0%; Score 37; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 346 HHPHG 350

RESULT 32
Q99PS7
ID Q99PS7 PRELIMINARY; PRT; 515 AA.
AC Q99PS7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein 2.
GN Name=RNHRG2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;

RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055896; BAB33093.1; --
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 515 AA; 58055 MW; 7CEBA3A1A3678966 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 346 HHPHG 350

RESULT 33
Q80XX4
ID Q80XX4 PRELIMINARY; PRT; 524 AA.
AC Q80XX4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Accn4 protein (Fragment).
GN Name=Accn4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the amiloride-sensitive sodium channel
CC family.
DR EMBL; BC046481; AAH46481.1; --
DR MGD; MGI:2652846; Accn4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005272; F:sodium channel activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001873; Na+channel_ASC.
DR Pfam; PF00858; ASC; 1.

DR PRINTS; PR01078; AMINACHANNEL.
KW Ion transport; Ionic channel; Sodium channel; Sodium transport;
KW Transmembrane; Transport.
FT NON TER 1
SQ SEQUENCE 524 AA; 57469 MW; AD1B1991F9B82FAF CRC64;

Query Match 100.0%; Score 37; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 509 HHPHG 513

RESULT 34
HRG_HUMAN
ID HRG_HUMAN STANDARD; PRT; 525 AA.
AC P04196;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein precursor (Histidine-proline rich
DE glycoprotein) (HPRG).
GN Name=HRG;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=86216149; PubMed=3011081;
RA Koide T., Foster D.C., Yoshitake S., Davie E.W.;
RT "Amino acid sequence of human histidine-rich glycoprotein derived from
RT the nucleotide sequence of its cDNA.";
RL Biochemistry 25:2220-2225 (1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 214-247 FROM N.A.
RX MEDLINE=94245171; PubMed=8188234;
RA Hennis B.C., Frants R.R., Bakker E., Vossen R.H., van der Poort E.W.,
RA Blonden L.A., Cox S., Khan P.M., Spurr N.K., Kluft C.;
RT "Evidence for the absence of intron H of the histidine-rich
RT glycoprotein (HRG) gene: genetic mapping and in situ localization of
RT HRG to chromosome 3q28-q29.";
RL Genomics 19:195-197 (1994).
RN [4]
RP SEQUENCE OF 19-27.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714 (1992).
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 12 tandem repetitions of a 5-
CC residue sequence (GHPH, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M13149; AAA52694.1; -.
DR EMBL; AB005803; BAA21613.1; -.
DR EMBL; Z17218; CAA78925.1; -.
DR PIR; A01287; KGHUGH.
DR SWISS-2DPAGE; P04196; HUMAN.
DR Genew; HGNC:5181; HRG.
DR MIM; 142640; -.
DR InterPro; IPR000010; Prot_inh_cystat.

DR Pfam; PF00031; Cystatin; 1.
KW Direct protein sequencing; Glycoprotein; Heparin-binding;
KW Polymorphism; Repeat; Signal.
FT SIGNAL 1 18

FT CHAIN 19 525 Histidine-rich glycoprotein.
FT DOMAIN 19 136 Cystatin-like 1.
FT DOMAIN 137 254 Cystatin-like 2.
FT DOMAIN 276 321 Pro-rich.
FT DOMAIN 350 497 His/Pro-rich.
FT DISULFID 24 504 By similarity.
FT DISULFID 78 89 By similarity.
FT DISULFID 105 126 By similarity.
FT DISULFID 203 417 By similarity.
FT DISULFID 218 241 By similarity.
FT CARBOHYD 63 63 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 87 87 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 125 125 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 344 344 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 345 345 N-linked (GlcNAc. . .) (Potential).
FT VARIANT 204 204 P -> S (in dbSNP:3181917).
FT /FTId=VAR_014528.

SQ SEQUENCE 525 AA; 59578 MW; A2B124D6CE93114F CRC64;

Query Match 100.0%; Score 37; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 384 HHPHG 388

RESULT 35
Q99PS5 ID Q99PS5 PRELIMINARY; PRT; 525 AA.
AC Q99PS5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg; Synonyms=MMHRG;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tsuchida N., Wakabayashi S., Jahnen-Dechent W., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AB055898; BAB33095.1; -.
DR EMBL; BC011168; AAH11168.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59162 MW; A83E93A439CFB126 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 361 HHPHG 365

RESULT 36
Q99PS6 ID Q99PS6 PRELIMINARY; PRT; 525 AA.
AC Q99PS6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg; Synonyms=MMHRG;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB055897; BAB33094.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59090 MW; A83E93A439CFB3AC CRC64;

Query Match 100.0%; Score 37; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 361 HHPHG 365

RESULT 37									
Q99PS8									
ID	Q99PS8	PRELIMINARY;			PRT;	525 AA.			
AC	Q99PS8;								
DT	01-JUN-2001	(TrEMBLrel. 17, Created)							
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)							
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)							
DE	Histidine-rich glycoprotein 1.								
GN	Name=RNHRG1;								
OS	Rattus norvegicus (Rat).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.								
OX	NCBI_TaxID=10116;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=Sprague-Dawley; TISSUE=Liver;								
RA	Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,								
RA	Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;								
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AB055895; BAB33092.1; -.								
DR	GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.								
DR	InterPro; IPR000010; Prot_inh_cystat.								
DR	Pfam; PF00031; Cystatin; I.								
DR	SMART; SM00043; CY; 2.								
SQ	SEQUENCE 525 AA; 59049 MW; 38290A631FAC7777 CRC64;								
Query Match 100.0%; Score 37; DB 2; Length 525;									
Best Local Similarity 100.0%; Pred. No. 2.2e+02;									
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1 HHPHG 5								
Db	346 HHPHG 350								
RESULT 38									
Q9ESB3									
ID	Q9ESB3	PRELIMINARY;			PRT;	525 AA.			
AC	Q9ESB3;								
DT	01-MAR-2001	(TrEMBLrel. 16, Created)							
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)							
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)							
DE	Histidine-rich glycoprotein.								
GN	Name=Hrg;								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxID=10090;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=129;								
RX	MEDLINE=20307726; PubMed=10849117;								
RA	Hulett M.D., Parish C.R.;								
RT	"Murine histidine-rich glycoprotein: cloning, characterization and								
RT	cellular origin.";								
RL	Immunol. Cell Biol. 78:280-287(2000).								
DR	EMBL; AF194028; AAG28416.1; -.								
DR	MGD; MGI:2146636; Hrg.								
DR	GO; GO:0005615; C:extracellular space; TAS.								
DR	InterPro; IPR000010; Prot_inh_cystat.								
DR	Pfam; PF00031; Cystatin; I.								
DR	SMART; SM00043; CY; 2.								
SQ	SEQUENCE 525 AA; 59132 MW; 6E55F2A439CFB123 CRC64;								
Query Match 100.0%; Score 37; DB 2; Length 525;									
Best Local Similarity 100.0%; Pred. No. 2.2e+02;									
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1 HHPHG 5								
Db	361 HHPHG 365								

RESULT 39

HRG_RABIT	STANDARD;	PRT;	526 AA.
ID	HRG_RABIT		
AC	Q28640;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG) (Fragment).		
GN	Name=HRG;		
OS	Oryctolagus cuniculus (Rabbit).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX	NCBI_TaxID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 9-23; 301-313 AND 422-429.		
RC	TISSUE=Serum;		
RX	MEDLINE=96229917; PubMed=8639676; DOI=10.1021/bi952061t;		
RA	Borza D.-B., Tatum F.M., Morgan W.T.;		
RT	"Domain structure and conformation of histidine-proline-rich glycoprotein.";		
RT	Biochemistry 35:1925-1934(1996).		
RL	-!- FUNCTION: The physiological function is not yet known. It binds heme, dyes and divalent metal ions. It can inhibit rosette formation and is known to interact with heparin, thrombospondin, and the lysine-binding site of plasminogen. On the basis of its homology with HMW kininogen, the His-rich region of this protein may mediate the contact activation phase of intrinsic blood coagulation cascade.		
CC	-!- SUBCELLULAR LOCATION: Secreted.		
CC	-!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.		
CC	-!- DOMAIN: In addition to having a high His and Pro content, this protein has many internal repeats. 15 tandem repetitions of a 5-residue sequence (G[H/P][H/P]PH, consensus) form a His/Pro-rich region.		
CC	-!- SIMILARITY: Contains 2 cystatin-like domains.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; U32189; AAC48516.1; -.		
DR	InterPro; IPR000010; Prot_inh_cystat.		
DR	Pfam; PF00031; Cystatin; I.		
KW	Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat;		
KW	Signal.		
FT	NON TER	1	1
FT	SIGNAL	<1	8
FT	CHAIN	9	526
FT	DOMAIN	9	126
FT	DOMAIN	127	243
FT	DOMAIN	251	296
FT	DOMAIN	329	498
FT	DISULFID	14	505
FT	DISULFID	68	79
FT	DISULFID	95	116
FT	DISULFID	193	415
FT	DISULFID	207	230
FT	DISULFID	272	302
FT	CARBOHYD	115	115
FT	CARBOHYD	192	192
FT	CARBOHYD	240	240
FT	CARBOHYD	310	310
FT	CARBOHYD	485	485
FT	SITE	303	304
FT	SITE	421	422
SQ	SEQUENCE	526 AA;	58877 MW; 810F23D367D93D42 CRC64;

Query Match 100.0%; Score 37; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|
|
|
|
Db 333 HHPHG 337

RESULT 40

Q6YK32 PRELIMINARY; PRT; 536 AA.
AC Q6YK32;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Histidine-rich glycoprotein HRG.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137504; AAN27996.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 536 AA; 60439 MW; CCF14FC08DD53D7C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 536;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|
|
|
|
Db 372 HHPHG 376

Search completed: June 15, 2005, 14:21:24
Job time : 85.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB Length	ID	Description
1	36	100.0	5	5	ABB79810	Abb79810 Histidine
2	36	100.0	5	8	ADH10417	Adh10417 Rabbit HP
3	36	100.0	20	4	ABB42309	Abb42309 Peptide #
4	36	100.0	20	4	AAM36117	Aam36117 Peptide #
5	36	100.0	20	4	ABB25810	Abb25810 Protein #
6	36	100.0	20	4	AAM76009	Aam76009 Human bon
7	36	100.0	20	4	AAM63195	Aam63195 Human bra
8	36	100.0	20	4	ABG57735	Abg57735 Human liv
9	36	100.0	20	5	ABG45414	Abg45414 Human pep
10	36	100.0	53	4	AAU46033	Aau46033 Propionib
11	36	100.0	53	6	ABM42552	Abm42552 Propionib
12	36	100.0	61	4	AAU54374	Aau54374 Propionib
13	36	100.0	61	6	ABM50893	Abm50893 Propionib
14	36	100.0	63	4	AAU22059	Aau22059 Human car
15	36	100.0	63	7	ADE46027	Ade46027 Human car
16	36	100.0	63	8	ADJ07445	Adj07445 Human car
17	36	100.0	87	4	AAU66176	Aau66176 Propionib
18	36	100.0	87	6	ABM62695	Abm62695 Propionib
19	36	100.0	92	4	AAU58530	Aau58530 Propionib
20	36	100.0	92	6	ABM55049	Abm55049 Propionib
21	36	100.0	94	4	AAU63724	Aau63724 Propionib
22	36	100.0	94	6	ABM60243	Abm60243 Propionib
23	36	100.0	99	4	AAU60014	Aau60014 Propionib
24	36	100.0	99	4	ABB15846	Abb15846 Human ner
25	36	100.0	99	6	ABM56533	Abm56533 Propionib

26	36	100.0	101	8	ADH10412	Adh10412 Rabbit HP
27	36	100.0	112	4	AAO13856	Aao13856 Human pol
28	36	100.0	117	5	ABG70306	Abg70306 Human MDD
29	36	100.0	118	4	AAU00910	Aau00910 Human can
30	36	100.0	202	4	AAU40328	Aau40328 Propionib
31	36	100.0	202	6	ABM36847	Abm36847 Propionib
32	36	100.0	231	2	AAU29941	Aay29941 Zea mays
33	36	100.0	241	6	ADA48350	Ada48350 Rice prot
34	36	100.0	262	4	AAU87400	Aau87400 Novel cen
35	36	100.0	262	7	ABO61857	Abo61857 Klebsiell
36	36	100.0	262	8	ADI54715	Adi54715 Novel hum
37	36	100.0	311	4	ABB66594	Abb66594 Drosophil
38	36	100.0	324	4	AAU51225	Aau51225 Propionib
39	36	100.0	324	6	ABM47744	Abm47744 Propionib
40	36	100.0	475	6	ABU33219	Abu33219 Protein e
41	36	100.0	480	7	ABO78053	Abo78053 Pseudomon
42	36	100.0	504	4	ABB68243	Abb68243 Drosophil
43	36	100.0	516	7	ABO71042	Abo71042 Pseudomon
44	36	100.0	526	5	ABB79805	Abb79805 Rabbit hi
45	36	100.0	526	8	ADH10410	Adh10410 Rabbit HP

ALIGNMENTS

RESULT 1
ABB79810
ID ABB79810 standard; peptide; 5 AA.

XX AC ABB79810;

XX DT 25-NOV-2002 (first entry)

XX DE Histidine proline rich glycoprotein pentapeptide.

XX KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW nootropic; neuroprotective; antiparkinsonian.

XX OS Synthetic.

XX PN WO200264621-A2.

XX PD 22-AUG-2002.

XX PF 14-FEB-2002; 2002WO-US004336.

XX PR 14-FEB-2001; 2001US-0268370P.

XX PA (ATTE-) ATTENUON LLC.

XX PI Donate F, Harris S, Plunkett ML, Mazar AP;

XX DR WPI; 2002-666989/71.

XX PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX PS Claim 2; Page 67; 82pp; English.

XX CC The present sequence is a specific example of claimed anti-angiogenic
CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
CC peptides comprise: the histidine-proline-rich (H/P) domain of human
CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808, such as the present peptide, or its variant
CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
CC N- or C-terminus. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or

CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 36; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 1 HPPHG 5

RESULT 2
ADH10417
ID ADH10417 standard; peptide; 5 AA.
XX
AC ADH10417;

XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein H/P rich domain repeat fragment.
XX

KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.

XX
OS Oryctolagus cuniculus.

XX
PN WO2003077872-A2.

XX
PD 25-SEP-2003.

XX
PF 17-MAR-2003; 2003WO-US008060.

XX
PR 15-MAR-2002; 2002US-0364047P.

XX
PA (ATTE-) ATTENUON LLC.

XX

PI Mcrae K, Donate F, Juarez J, Mazar AP;
XX
DR WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Example 5; SEQ ID NO 31; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC an antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 36; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 1 HPPHG 5

RESULT 3
ABB42309
ID ABB42309 standard; peptide; 20 AA.
XX
AC ABB42309;

XX
DT 04-FEB-2002 (first entry)

XX
DE Peptide #9815 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX
OS Homo sapiens.

XX
PN WO200157277-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US000669.

XX
PR 04-FEB-2000; 2000US-0180312P.

XX
PR 26-MAY-2000; 2000US-0207456P.

XX
PR 30-JUN-2000; 2000US-00608408.

XX
PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX Claim 27; SEQ ID NO 34944; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 20 AA;
SQ Query Match 100.0%; Score 36; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HPPHG 5
Db |||||
3 HPPHG 7
RESULT 4
ID AAM36117 standard; protein; 20 AA.
XX
AC AAM36117;
XX 17-OCT-2001 (first entry)
DT Peptide #10154 encoded by probe for measuring placental gene expression.
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX Homo sapiens.
XX WO200157272-A2.
PN 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US0000663.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX

PS Claim 27; SEQ ID NO 36386; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SNP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX Sequence 20 AA;
SQ Query Match 100.0%; Score 36; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HPPHG 5
Db |||||
3 HPPHG 7
RESULT 5
ID ABB25810 standard; protein; 20 AA.
XX
AC ABB25810;
XX 23-JAN-2002 (first entry)
DT Protein #7809 encoded by probe for measuring heart cell gene expression.
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX Homo sapiens.
XX WO200157274-A2.
PN 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US0000666.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX Claim 15; SEQ ID NO 27580; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

```
XX SQ      Sequence 20 AA;
      Query Match      100.0%; Score 36; DB 4; Length 20;
      Best Local Similarity 100.0%; Pred. No. 34;
      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      3 HPPHG 7

RESULT 6
AAM76009
ID      AAM76009 standard; protein; 20 AA.
XX
AC      AAM76009;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed probe encoded protein SEQ ID NO: 36315.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US000668.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488900/53.
XX
PF      30-JAN-2001; 2001WO-US000668.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488900/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for analyzing
PT      gene expression in human bone marrow.
XX
PS      Example 4; SEQ ID NO 36315; 658pp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      bone marrow. They can be used to measure gene expression in bone marrow
CC      samples, which may enable the improved diagnosis and treatment of cancers
CC      such as lymphoma, leukaemia and myeloma. The present sequence is a
CC      protein encoded by one of the probes of the invention
XX
SQ      Sequence 20 AA;

      Query Match      100.0%; Score 36; DB 4; Length 20;
      Best Local Similarity 100.0%; Pred. No. 34;
      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      3 HPPHG 7

RESULT 7
AAM63195
ID      AAM63195 standard; protein; 20 AA.
```

```
XX AAM63195;
XX
DT      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe encoded protein SEQ ID NO: 35300.
XX
KW      Human; brain expressed exon; gene expression analysis; probe; microarray;
KW      Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS      Homo sapiens.
XX
PN      WO200157275-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US000667.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483446/52.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      brains.
XX
PS      Example 4; SEQ ID NO 35300; 650pp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      brain. They can be used to measure gene expression in brain cell samples,
CC      which may enable the diagnosis and improved treatment of nervous system
CC      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of
CC      epilepsy and cancers. The present sequence is a protein encoded by one of
CC      the probes of the invention
XX
SQ      Sequence 20 AA;

      Query Match      100.0%; Score 36; DB 4; Length 20;
      Best Local Similarity 100.0%; Pred. No. 34;
      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      3 HPPHG 7

RESULT 8
ABG57735
ID      ABG57735 standard; peptide; 20 AA.
XX
AC      ABG57735;
XX
DT      25-FEB-2003 (first entry)
XX
DE      Human liver peptide, SEQ ID No 36383.
XX
KW      Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW      hypercholesterolaemia; coronary heart disease.
XX
OS      Homo sapiens.
XX
PN      WO200157273-A2.
XX
```

PD	09-AUG-2001.	PD	15-NOV-2001.
XX		XX	
PF	30-JAN-2001; 2001WO-US000664.	PF	30-JAN-2001; 2001WO-US000665.
XX		XX	
PR	04-FEB-2000; 2000US-0180312P.	PR	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.	PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.	PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.	PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.	PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.	PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.	PR	04-OCT-2000; 2000GB-00024263.
XX		XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.	PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX		XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX		XX	
DR	WPI; 2001-488898/53.	DR	WPI; 2002-114183/15.
XX		XX	
PT	Human genome-derived single exon nucleic acid probes useful for analyzing	PT	Spatially-addressable set of single exon nucleic acid probes, used to
PT	gene expression in human adult liver.	PT	measure gene expression in human lung samples.
XX		XX	
PS	Claim 27; SEQ ID NO 36383; 658pp; English.	PS	Claim 27; SEQ ID NO 35079; 634pp; English.
XX		XX	
CC	The invention relates to a single exon nucleic acid probe (SENP) (I) for	CC	The invention relates to a spatially-addressable set of single exon
CC	measuring human gene expression in a sample derived from human adult	CC	nucleic acid probes for measuring gene expression in a sample derived
CC	liver, comprising one of 13109 defined nucleotide sequences given in the	CC	from human lung comprising single exon nucleic acid probes having one of
CC	specification (or complements/ fragments). The probe hybridises at high	CC	12614 nucleic acid sequences mentioned in the specification, or their
CC	stringency to a nucleic acid molecule expressed in the human adult liver.	CC	complements or the 12387 open reading frames derived from the 12614
CC	(I) may be used for predicting, measuring and displaying gene expression	CC	probes. Also included are a microarray comprising the novel set of probes
CC	in samples derived from human adult liver. The genes identified may be	CC	; the novel set of probes which hybridise at high stringency to a nucleic
CC	involved in genetic liver diseases such as cirrhosis,	CC	acid expressed in the human lung; measuring gene expression in a sample
CC	hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is	CC	derived from human lung, comprising (a) contacting the array with a
CC	associated with coronary heart disease. ABG47348-ABG59930 represent human	CC	collection of detectably labeled nucleic acids derived from human lung
CC	liver single exon encoded peptides of the invention. Note: The sequence	CC	mRNA, and (b) measuring the label detectably bound to each probe of the
CC	information for this patent does not appear in the printed specification	CC	array; identifying exons in a eukaryotic genome, comprising (a)
CC	but was obtained in electronic format directly from WIPO at	CC	algorithmically predicting at least one exon from genomic sequences of
CC	ftp.wipo.int/pub/published_pct_sequences	CC	the eukaryote; and (b) detecting specific hybridisation of detectably
XX		CC	labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
SQ	Sequence 20 AA;	CC	having a fragment identical to the predicted exon, the probe is included
		CC	in the above mentioned microarray; assigning exons to a single gene,
		CC	comprising (a) identifying exons from genomic sequence by the method
		CC	above and (b) measuring the expression of each of the exons in several
		CC	tissues and/or cell types using hybridisation to a single exon
QY	1 HPPHG 5	CC	microarrays having a probe with the exon, where a common pattern of
Db		CC	expression of the exons in the tissues and/or cell types indicates that
	3 HPPHG 7	CC	the exons should be assigned to a single gene; a peptide comprising one
		CC	of 12011 sequences, mentioned in the specification, or encoded by the
RESULT 9		CC	probes/open reading frames (ORF). The probes are used for gene expression
ABG45414		CC	analysis, and for identifying exons in a gene, particularly using human
ID	ABG45414 standard; peptide; 20 AA.	CC	lung derived mRNA and for the study of lung diseases such as asthma, lung
XX		CC	cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
AC	ABG45414;	CC	disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX		CC	tuberosus sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
DT	19-AUG-2002 (first entry)	CC	Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX		CC	histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
DE	Human peptide encoded by genome-derived single exon probe SEQ ID 35079.	CC	Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX		CC	dyskinesia, pulmonary hypertension and hyaline membrane disease. The
KW	Human; single exon probe; asthma; lung cancer; COPD; ILD;	CC	present sequence is a peptide/protein encoded by a single exon probe of
KW	chronic obstructive pulmonary disease; interstitial lung disease;	CC	the invention. Note: The sequence data for this patent did not form part
KW	familial idiopathic pulmonary fibrosis; neurofibromatosis;	CC	of the printed specification, but was obtained in electronic format
KW	tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;	CC	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
KW	Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;	XX	
KW	pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;	SQ	Sequence 20 AA;
KW	pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;		
KW	primary ciliary dyskinesia; pulmonary hypertension;		
KW	hyaline membrane disease.		
OS	Homo sapiens.	QY	1 HPPHG 5
XX		Db	
XX			3 HPPHG 7
PN	WO200186003-A2.		
XX			

Query Match 100.0%; Score 36; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 3 HPPHG 7

RESULT 10
AAU46033
ID AAU46033 standard; protein; 53 AA.
XX
AC AAU46033;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #6929.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59529.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 7228; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 36; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
 |||||
Db 7 HPPHG 11

RESULT 11
ABM42552
ID ABM42552 standard; protein; 53 AA.
XX
AC ABM42552;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #7228.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64458.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 7228; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 36; DB 6; Length 53;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

Db		7 HPPHG 11	Db		12 HPPHG 16
RESULT 12					
AAU54374					
ID	AAU54374	standard; protein; 61 AA.	ID	ABM50893	standard; protein; 61 AA.
XX	AC	AAU54374;	XX	AC	ABM50893;
XX	DT	27-FEB-2002 (first entry)	XX	DT	20-OCT-2003 (first entry)
XX	DE	Propionibacterium acnes immunogenic protein #15270.	XX	DE	Propionibacterium acnes predicted ORF-encoded polypeptide #15569.
XX	KW	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;	XX	KW	Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW	KW	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;	KW	KW	immunostimulant; immune response; vaccine.
KW	KW	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;	XX	OS	Propionibacterium acnes.
XX	XX	dermatological; osteopathic; neuroprotectant.	XX	PN	WO2003033515-A1.
OS	XX	Propionibacterium acnes.	XX	PN	WO2003033515-A1.
XX	PN	WO200181581-A2.	XX	PD	24-APR-2003.
XX	PD	01-NOV-2001.	XX	PF	11-OCT-2002; 2002WO-US032727.
XX	PF	20-APR-2001; 2001WO-US012865.	XX	PR	15-OCT-2001; 2001US-00978825.
XX	PR	21-APR-2000; 2000US-0199047P.	XX	PA	(CORI-) CORIXA CORP.
PR	PR	02-JUN-2000; 2000US-0208841P.	XX	PI	Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PR	PR	07-JUL-2000; 2000US-0216747P.	PI	PI	Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX	XX	(CORI-) CORIXA CORP.	PI	PI	Barth B, Vallieve-Douglass J;
PA	XX		XX	XX	
XX	PI	Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;	DR	WPI	WPI; 2003-381789/36.
PI	PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;	DR	N-PSDB	N-PSDB; ACF64493.
XX	XX		XX	XX	
DR	DR	WPI; 2001-616774/71.	PT	PT	New Propionibacterium acnes polypeptides and polynucleotides encoding the
DR	DR	N-PSDB; AAS59564.	PT	PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX	XX		PT	XX	or for stimulating an immune response specific for a P. acnes protein.
PT	PT	Propionibacterium acnes polypeptides and nucleic acids useful for	PS	PS	Example 1; SEQ ID NO 15569; 1481pp; English.
PT	PT	vaccinating against and diagnosing infections, especially useful for	XX	XX	
PT	PT	treating acne vulgaris.	XX	XX	
XX	XX		CC	CC	The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX	PS	Example 1; SEQ ID NO 15569; 1069pp; English.	CC	CC	encoding a Propionibacterium acnes protein. The invention also relates to
XX	XX		CC	CC	polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC	CC	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic	CC	CC	immunogenic fragments of P. acnes polypeptides. The invention
CC	CC	polypeptides. The proteins and their associated DNA sequences are used in	CC	CC	additionally encompasses expression vectors and host cells comprising a
CC	CC	the treatment, prevention and diagnosis of medical conditions caused by	CC	CC	polynucleotide of the invention; antibodies against polypeptides of the
CC	CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,	CC	CC	invention; fusion proteins comprising a polypeptide of the invention; a
CC	CC	pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.	CC	CC	method for stimulating an immune response specific for a P. acnes
CC	CC	P. acnes is also involved in infections of bone, joints and the central	CC	CC	polypeptide and an isolated T cell population comprising T cells prepared
CC	CC	nervous system, however it is particularly involved in the inflammatory	CC	CC	via this method; a vaccine composition (comprising P. acnes polypeptides,
CC	CC	lesions associated with acne vulgaris. A method for detecting the	CC	CC	polynucleotides, antibodies, fusion proteins, T cell populations, or
CC	CC	presence or absence of P. acnes in a patient comprises contacting a	CC	CC	antigen-presenting cells that express the polypeptide); a method and kit
CC	CC	sample with a binding agent that binds to the proteins of the invention	CC	CC	for detecting or determining the presence or absence of P. acnes in a
CC	CC	and determining the amount of bound protein in the sample. The	CC	CC	patient; and a method for inhibiting the development of P. acnes in a
CC	CC	polypeptides may be used as antigens in the production of antibodies	CC	CC	proteins. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC	CC	specific for P. acnes proteins. These antibodies can be used to	CC	CC	polypeptides are useful for diagnosing, preventing or treating acne
CC	CC	downregulate expression and activity of P. acnes polypeptides and	CC	CC	vulgaris, or for stimulating an immune response specific for a P. acnes
CC	CC	therefore treat P. acnes infections. The antibodies may also be used as	CC	CC	protein. The polynucleotides can also be used as probes or primers for
CC	CC	diagnostic agents for determining P. acnes presence, for example, by	CC	CC	nucleic acid hybridisation. The vaccine composition is useful for the
CC	CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for	CC	CC	stimulation of an immune response against P. acnes, or for treating acne,
CC	CC	this patent did not form part of the printed specification, but was	CC	CC	and the kit is useful for performing a diagnostic assay. The present
CC	CC	obtained in electronic format directly from WIPO at	CC	CC	sequence represents a polypeptide predicted to be encoded by an ORF (open
CC	CC	ftp.wipo.int/pub/published_pct_sequences	CC	CC	reading frame) contained within the P. acnes polynucleotides of the
XX	XX		CC	CC	invention. Note: The sequence data for this patent did not form part of
SQ	SQ	Sequence 61 AA;	CC	CC	the printed specification, but was obtained in electronic format directly
Query Match	Query Match	100.0%; Score 36; DB 4; Length 61;	CC	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
Best Local Similarity	Best Local Similarity	100.0%; Pred. No. 96;	XX	XX	
Matches	Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	SQ	SQ	Sequence 61 AA;
Qy	Qy	1 HPPHG 5	Query Match	Query Match	100.0%; Score 36; DB 6; Length 61;

Best Local Similarity 100.0%; Pred. No. 96;			
Matches	5; Conservative	0; Mismatches	0; Indels
			0; Gaps
QY	1 HPPHG 5		
Db	12 HPPHG 16		
RESULT 14			
AAU22059			
ID	AAU22059 standard; protein; 63 AA.		
XX			
AC	AAU22059;		
XX			
DT	18-DEC-2001 (first entry)		
XX			
DE	Human cardiovascular system antigen polypeptide SEQ ID No 833.		
XX			
KW	Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;		
KW	chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;		
KW	antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;		
KW	cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;		
KW	ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;		
KW	hyperproliferative disorder; breast; liver; cardiovascular disorder;		
KW	cerebrovascular disorder; nervous system disorder; bacterial infection;		
KW	fungal infection; viral infection; ocular disorder; endocrine disorder;		
KW	gastrointestinal disorder; renal disorder; respiratory disorder;		
KW	wound healing; skin aging; organ transplantation; tissue regeneration;		
KW	anti-infertility.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200155321-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001340.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	18-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226868P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		

PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451930/48.

N-PSDB; AAS35333.

New cardiovascular system related polynucleotides and polypeptides, useful for diagnosing, treating and/or preventing disorders of the cardiovascular system.

Claim 11; SEQ ID NO 833; 674pp; English.

Sequences AAU21852-AAU22466 represent the cardiovascular system antigen polypeptides of the invention. Cardiovascular system antigens and their associated polynucleotides are useful in the diagnosis, treatment and prevention of various types of disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological condition can be determined by detecting the presence or absence of a mutation in a cardiovascular system antigen polynucleotide. The treatable disorders include autoimmune diseases such as rheumatoid arthritis, hyperproliferative disorders such as neoplasms of the breast or liver, cardiovascular disorders such as cardiac arrest, cerebrovascular disorders such as cerebral ischaemia, nervous system disorders such as Alzheimer's disease, infections caused by bacteria, viruses and fungi, ocular disorders such as corneal infection, endocrine disorders such as premature labour and infertility, gastrointestinal disorders such as Crohn's disease, renal disorders such as glomerulonephritis and respiratory disorders such as asthma and pleurisy. The polypeptides can also be used to aid wound healing, to prevent skin aging due to sunburn, to maintain organs before transplantation, to regenerate tissues and in chemotaxis. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 100.0%; Score 36; DB 4; Length 63;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5

|||||

26 HPPHG 30

Db

RESULT 15
ADE46027
ID ADE46027 standard; protein; 63 AA.
XX
AC ADE46027;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human cardiovascular system related polypeptide #208.
XX
KW Human; cardiovascular system related polypeptide; cancer;
KW proliferative disorder; foetal abnormality; developmental abnormality;
KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder.
XX
OS Homo sapiens.
XX
PN US2003059908-A1.
XX
PD 27-MAR-2003.
XX
PF 07-MAR-2002; 2002US-00091504.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-743766/70.
DR N-PSDB; ADE45412.
XX
PT New cardiovascular system related polynucleotides and polypeptides,
PT useful for preventing, treating, or ameliorating a medical condition,
PT such as cancer of cardiovascular tissues and cancer metastases.
XX
PS Claim 11; SEQ ID NO 833; 262pp; English.
XX
CC The invention relates to human cardiovascular system related polypeptides
CC and the polynucleotides encoding them. The polypeptides, polynucleotides
CC and antibodies to the polypeptides are useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition,
CC for preventing, treating, or ameliorating a medical condition, such as
CC cancer of cardiovascular system tissues, proliferative disorders, foetal
CC and developmental abnormalities, haematopoietic disorders, diseases of
CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
CC arthritis), inflammation, allergies, neurological disorders (e.g.,
CC Alzheimer's disease, Parkinson's disease), cognitive disorders,
CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,
CC diabetes, atherosclerosis, cardiovascular disorders, angiogenic
CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-
CC related disorders, endocrine disorders and infections. The nucleic acids
CC are also useful for chromosome identification, radiation hybrid mapping
CC or long-range restriction mapping. The polypeptides and polynucleotides
CC may also be used as food additives or preservatives to increase or
CC decrease storage capabilities, fat content or other nutritional
CC components. This sequence represents a human cardiovascular system
CC related polypeptide of the invention.
XX
SQ Sequence 63 AA;

Query Match 100.0%; Score 36; DB 7; Length 63;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 26 HPPHG 30

RESULT 16
ADJ07445
ID ADJ07445 standard; protein; 63 AA.
XX
AC ADJ07445;
XX
DT 04-NOV-2004 (first entry)

XX Human cardiovascular system associated polypeptide SeqID833.
DE
XX
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW breast neoplasms; liver neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn;
KW organ transplantation; cell culture; tissue regeneration; chemotaxis;
KW food additive; preservative; cardiovascular system associated antigen;
KW nuclear factor kappaB; NFkappaB; promoter element; human.
XX
OS Homo sapiens.
XX
PN US2004005575-A1.
XX
PD 08-JAN-2004.
XX
PF 26-AUG-2002; 2002US-00227577.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
PR 07-MAR-2002; 2002US-00091504.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2004-081713/08.
DR N-PSDB; ADJ06830.
XX
PT New cardiovascular system-related nucleic acid molecule, useful for
PT diagnosing, preventing or treating diseases of the cardiovascular system,
PT and in chromosome mapping, drug screening or in pharmacogenomics.
XX
PS Claim 11; SEQ ID NO 833; 262pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC human cardiovascular system associated polypeptide (or antigens), or its
CC fragment. Also included recombinant vectors, recombinant host cells, an
CC isolated human cardiovascular system associated polypeptide (including
CC its fragment, allelic variant, species homologue or epitope), an isolated
CC antibody that binds specifically to a human cardiovascular system
CC associated polypeptide, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in human cardiovascular system
CC associated nucleic acid and diagnosing a condition based on the presence
CC or absence of the mutation), identifying a binding partner to human
CC cardiovascular system associated polypeptides, the gene corresponding to
CC the human cardiovascular system associated cDNA sequence and identifying
CC an activity in a biological assay comprising expressing the human
CC cardiovascular system associated cDNA in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity. The human
CC cardiovascular system associated nucleic acids and polypeptides are used
CC to prevent, treat or ameliorate a medical condition (for example in
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for
CC example autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders, for example neoplasms of the breast or

Query Match 100.0%; Score 36; DB 8; Length 63;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
|||||
Db 26 HPPHG 30

RESULT 17
AAU66176
ID AAU66176 standard; protein; 87 AA.
XX
AC AAU66176;
XX

DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #27072.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59717.
XX

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 27371; 1069pp; English.

XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 87 AA;

Query Match 100.0%; Score 36; DB 4; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
|||||
Db 57 HPPHG 61

RESULT 18
ABM62695
ID ABM62695 standard; protein; 87 AA.
XX
AC ABM62695;
XX

DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #27371.

XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.

XX WO2003033515-A1.
PN 24-APR-2003.
XX 11-OCT-2002; 2002WO-US032727.
XX 15-OCT-2001; 2001US-00978825.
PR (CORI-) CORIXA CORP.
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglass J;
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64646.
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 27371; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 36; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
57 HPPHG 61

RESULT 19
AAU58530
ID AAU58530 standard; protein; 92 AA.
XX
AC AAU58530;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #19426.
XX

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
PF
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59592.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 19725; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 92 AA;

Query Match 100.0%; Score 36; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
77 HPPHG 81

RESULT 20
ABM55049
ID ABM55049 standard; protein; 92 AA.
XX
AC ABM55049;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #19725.
XX

KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieue-Douglass J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64521.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 19725; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
CC via this method; a vaccine composition (comprising P. acnes polypeptides, or
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 92 AA;

Query Match 100.0%; Score 36; DB 6; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

Db |||||
77 HPPHG 81

RESULT 21

AAU63724

ID AAU63724 standard; protein; 94 AA.

XX

AC AAU63724;

XX

DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #24620.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59634.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 24919; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, uveitis,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 94 AA;

Query Match 100.0%; Score 36; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

Db |||||
64 HPPHG 68

RESULT 22

ABM60243

ID ABM60243 standard; protein; 94 AA.

XX

AC ABM60243;

XX

DT 20-OCT-2003 (first entry)
XX Propionibacterium acnes predicted ORF-encoded polypeptide #24919.
DE
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
XX WO2003033515-A1.
PN
XX
PD 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
PF
XX 15-OCT-2001; 2001US-00978825.
PR
XX (CORI-) CORIXA CORP.
PA
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64563.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 24919; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 94 AA;

Query Match 100.0%; Score 36; DB 6; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
|||||
Db 64 HPPHG 68

RESULT 23
AAU60014

ID AAU60014 standard; protein; 99 AA.
XX
AC AAU60014;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #20910.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
PA
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59607.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 21209; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 99 AA;

Query Match 100.0%; Score 36; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
|||||
Db 61 HPPHG 65

RESULT 24
ABB15846

ID ABB15846 standard; protein; 99 AA.
XX
AC ABB15846;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polypeptide SEQ ID NO 4503.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
DR WPI; 2001-541565/60.
DR N-PSDB; ABA12172.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases.
PT
PT
XX
PS Claim 11; SEQ ID NO 4503; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins (ABB14678-ABB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 99 AA;
Query Match 100.0%; Score 36; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 23 HPPHG 27
RESULT 25
ABM56533
ID ABM56533 standard; protein; 99 AA.
XX
AC ABM56533;
XX
XX 20-OCT-2003 (first entry)
DT
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #21209.
DE
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
XX WO2003033515-A1.

XX 24-APR-2003.
PD
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64536.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the polypeptide, useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein.
PS Example 1; SEQ ID NO 21209; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 99 AA;
Query Match 100.0%; Score 36; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 61 HPPHG 65
RESULT 26
ADH10412
ID ADH10412 standard; protein; 101 AA.
XX
XX ADH10412;
XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein H/P rich domain.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;

KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
XX angiogenesis; gene therapy; rabbit.
OS Oryctolagus cuniculus.
XX WO2003077872-A2.
PN
XX 25-SEP-2003.
PD
XX 17-MAR-2003; 2003WO-US008060.
PF
XX 15-MAR-2002; 2002US-0364047P.
PR
XX (ATTE-) ATTENUON LLC.
PA
XX Mccrae K, Donate F, Juarez J, Mazar AP;
PI WPI; 2004-090604/09.
XX
DR
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Claim 6; SEQ ID NO 24; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC an antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 36; DB 8; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
33 HPPHG 37

RESULT 27
AAO13856
ID AAO13856 standard; protein; 112 AA.
XX
AC AAO13856;
XX
DT 06-NOV-2001 (first entry)
XX

DE Human polypeptide SEQ ID NO 27748.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
PF
XX 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-S14838/56.
XX N-PSDB; AAI93787.
DR
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX Claim 20; SEQ ID NO 27748; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 112 AA;

Query Match 100.0%; Score 36; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
96 HPPHG 100

RESULT 28
ABG70306
ID ABG70306 standard; protein; 117 AA.
XX
AC ABG70306;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human MDDT protein Incyte ID No: LI:180252.16.orf2:2001JAN12.
XX
KW Human; molecule for disease detection and treatment; MDDT; cancer;
KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
KW autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antiasthmatic.
XX
OS Homo sapiens.
XX

PN WO200255738-A2.
XX 18-JUL-2002.
PD 09-JAN-2002; 2002WO-US001008.
XX 12-JAN-2001; 2001US-0261622P.
PR 16-JAN-2001; 2001US-0261865P.
PR 17-JAN-2001; 2001US-0262208P.
PR 17-JAN-2001; 2001US-0262209P.
PR 17-JAN-2001; 2001US-0262326P.
PR 19-JAN-2001; 2001US-0263063P.
PR 19-JAN-2001; 2001US-0263065P.
PR 19-JAN-2001; 2001US-0263329P.
XX (INCY-) INCYTE GENOMICS INC.
XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
PI Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ;
PI Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA;
XX WPI; 2002-590679/63.
DR N-PSDB; ABS51779.
XX New disease detection and treatment molecule (MDDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDDT expression, e.g. autoimmune or inflammatory
PT disorders.
XX Claim 27; Page 107; 129pp; English.
PS The present invention relates to the isolation of novel human molecules
XX for disease detection and treatment (MDDT), and the polynucleotide
CC sequences (mddt) encoding them. The MDDT polypeptides may be used to
CC screen for molecules that bind to, or are bound by the encoded
CC polypeptides, and to develop a transcript image of a tissue or cell type.
CC Probes comprising at least 20 nucleotides of the mddt polynucleotide may
CC be used to assess the toxicity of a test compound. The MDDT polypeptides
CC and mddt polynucleotides are useful in the diagnosis, study, prevention
CC and treatment of diseases associated with the expression of molecules for
CC disease detection and treatment. Such disorders include cell
CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, or cancers),
CC and autoimmune/inflammatory disorders (e.g. asthma, Crohn's disease, or
CC multiple sclerosis). The mddt polynucleotides may also be used as
CC molecule markers, in microarrays, and in somatic or germline gene
CC therapy. ABG70306-ABG70341 represent the MDDT proteins of the invention
XX Sequence 117 AA;
SQ Query Match 100.0%; Score 36; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 17 HPPHG 21
RESULT 29
AAU00910
ID AAU00910 standard; protein; 118 AA.
XX AAU00910;
AC AAU00910;
XX 04-JUL-2001 (first entry)
XX Human cancer related protein 25#1.
XX Human; cancer related protein; food additive; preservative; immunogen;
KW antibody; bone cancer; adrenal cancer; bone marrow cancer; breast cancer;
KW gastrointestinal cancer; liver cancer; lung cancer; urogenital cancer;
KW immune disorder; Addison's disease; allergy; diabetes mellitus;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; Crohn's disease;

KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW acquired immunodeficiency syndrome; AIDS; cardiovascular disorder;
KW myocardial ischaemia; wound healing; neurological disorder;
KW Parkinson's disease; Alzheimer's disease; cerebral anoxia; epilepsy;
KW viral infection; bacterial infection; fungal infection;
KW parasitic infection; agonist; antagonist.
XX Homo sapiens.
OS WO200118014-A1.
XX 15-MAR-2001.
PN 30-AUG-2000; 2000WO-US023794.
PD 03-SEP-1999; 99US-0152296P.
XX 06-OCT-1999; 99US-0158003P.
PR (HUMA-) HUMAN GENOME SCI INC.
PA Roschke V;
XX WPI; 2001-235186/24.
DR N-PSDB; AAS00851.
XX Twenty nine nucleic acid molecules encoding human cancer associated
PT proteins, useful in the prevention, treatment and diagnosis of cancer,
PT immune disorders, cardiovascular disorders and neurological diseases.
XX Disclosure; Page 422-423; 427pp; English.
PS The sequence represents a novel Human cancer related protein. The
XX polynucleotides and polypeptides are useful for preventing, treating or
CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. The polynucleotide are useful for chromosome
CC identification. The nucleic acids, protein, antibodies, agonists and
CC antagonists are useful in the diagnosis, treatment and prevention of
CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
CC myocardial ischaemias, wound healing, neurological diseases (e.g.
CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
CC and infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Numerous examples of each type of disorder are given in the
CC specification
XX Sequence 118 AA;
SQ Query Match 100.0%; Score 36; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 96 HPPHG 100
RESULT 30
AAU40328
ID AAU40328 standard; protein; 202 AA.
XX AAU40328;
AC AAU40328;
XX 13-FEB-2002 (first entry)
DT Propionibacterium acnes immunogenic protein #1224.
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD
XX 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59511.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 1523; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 202 AA;
Query Match 100.0%; Score 36; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 93 HPPHG 97
RESULT 31
ABM36847
ID ABM36847 standard; protein; 202 AA.
XX
AC ABM36847;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #1523.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64440.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 1523; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 202 AA;
Query Match 100.0%; Score 36; DB 6; Length 202;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 93 HPPHG 97
RESULT 32
AAY29941
ID AAY29941 standard; protein; 231 AA.
XX
AC AAY29941;
XX
DT 22-NOV-1999 (first entry)

XX Zea mays pathogenesis-related class I PR-1#70 protein.
DE
XX
KW Zea mays; maize; pathogenesis-related class I; PR-1; promoter;
KW regulation; expression; disease resistance; genetic manipulation;
KW tobacco mosaic virus; cucumber mosaic virus; ringspot virus;
KW necrosis virus; maize dwarf virus; viroid; bacterial; insect; nematode;
KW fungal.
XX
OS Zea mays.
XX
PN WO9943819-A1.
XX
PD 02-SEP-1999.
XX
PF 11-FEB-1999; 99WO-US003011.
XX
PR 26-FEB-1998; 98US-0076100P.
PR 27-MAR-1998; 98US-0079648P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX Crane VC;
XX
DR WPI; 1999-527621/44.
DR N-PSDB; AAZ21192.
XX
PT New promoter sequences from pathogenesis-related genes of maize.
XX
XX Claim 22; Page 69-70; 86pp; English.
XX
CC AAZ21186 to AAZ21190 represents the nucleotide sequences for promoters
CC isolated from a family of maize (Zea mays) genes encoding pathogenesis
CC related (PR-1) proteins. The promoters are useful for expressing
CC heterologous genes (including genes for disease resistance) in plants,
CC especially dicots, or monocots i.e. maize. The promoters are useful for
CC the genetic manipulation of plants to exhibit specific phenotypes,
CC particularly enhanced resistance to pathogen-caused disease. Pathogens
CC include viruses such as tobacco or cucumber mosaic virus, ringspot virus,
CC necrosis virus, and maize dwarf virus, and viroids, bacteria, insects,
CC nematodes and fungi. The present sequence represents a maize PR-1 protein
CC given in the present invention
XX
SQ Sequence 231 AA;

Query Match 100.0%; Score 36; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 174 HPPHG 178

RESULT 33
ADA48350
ID ADA48350 standard; protein; 241 AA.
XX
AC ADA48350;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice protein conferring disease resistance in plants.
XX
KW disease resistance; pathogen tolerance; plant pathogen; plant; rice.
XX
OS Oryza sativa.
XX
PN WO2003000906-A2.
XX
PD 03-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-IB002453.

XX 22-JUN-2001; 2001US-0300112P.
PR 26-SEP-2001; 2001US-0352277P.
PR 22-MAR-2002; 2002US-0366535P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Katagiri F, Krepes J, Provart N, Ricke D, Zhu T;
PI
XX
DR WPI; 2003-184052/18.
DR N-PSDB; ADA48349.
XX
PT New polynucleotide comprising a plant nucleotide sequence having an open
PT reading frame that encodes a polypeptide associated with disease
PT resistance, useful for conferring resistance or tolerance to a plant
PT pathogen.
XX
PS Claim 10; SEQ ID NO 420; 299pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a
CC plant nucleotide sequence having an open reading frame that encodes a
CC polypeptide associated with disease resistance or its fragment having
CC substantially the same activity as the full-length polypeptide. The
CC polynucleotide of the invention is useful for conferring resistance or
CC tolerance to a plant pathogen. The present sequence represents a protein
CC conferring disease resistance used in the invention.
XX
SQ Sequence 241 AA;

Query Match 100.0%; Score 36; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 178 HPPHG 182

RESULT 34
AAU87400
ID AAU87400 standard; protein; 262 AA.
XX
AC AAU87400;
XX
DT 05-JUN-2002 (first entry)
XX
DE Novel central nervous system protein #310.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155318-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001332.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-581633/65.

N-PSDB; ABK43730.

New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 9; SEQ ID NO 918; 837pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding a

CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 100.0%; Score 36; DB 4; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
17 HPPHG 21

RESULT 35
ABO61857
ID ABO61857 standard; protein; 262 AA.

XX ABO61857;
XX 29-JUL-2004 (first entry)
XX Klebsiella pneumoniae polypeptide seqid 8374.

KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.
XX US6610836-B1.
XX 26-AUG-2003.
XX 27-JAN-2000; 2000US-00489039.
XX 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.

PI Breton GL, Osborne M;
XX WPI; 2003-895346/82.
DR N-PSDB; ACH95408.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 8374; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a

CC Klebsiella pneumoniae polypeptide of the invention
XX Sequence 262 AA;

Query Match 100.0%; Score 36; DB 7; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
18 HPPHG 22

RESULT 36
ADIS4715
ID ADIS4715 standard; protein; 262 AA.

XX ADIS4715;
AC ADIS4715;
XX 22-APR-2004 (first entry)
XX Novel human protein seq id 918.

XX neuroprotective; nootropic; antiparkinsonian; anticonvulsant;
KW antidiabetic; antirheumatic; antiarthritic; dermatological;
KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiac;
KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
KW amyotrophic lateral sclerosis; multiple sclerosis;
KW immune system disorder; diabetes; rheumatoid arthritis;
KW systemic lupus erythematosus; autoimmune thyroiditis;
KW inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; infectious disease;
KW HIV infection; hepatitis infection; bacterial infection;
KW fungal infection; parasitic infection; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
KW renal disorder; acute glomerulonephritis; pyelonephritis;
KW renal lithiasis; proliferative disorder; cancerous diseases; human.

XX Homo sapiens.

XX US2004018969-A1.

XX 29-JAN-2004.

XX 17-JAN-2001; 2001US-00764875.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

DR WPI; 2004-122079/12.

DR N-PSDB; ADI54117.

XX
PT New polypeptides and nucleic acid molecules, useful for detecting,
PT preventing, diagnosing, prognosticating, treating or ameliorating medical
PT conditions e.g. neural disorders, reproductive disorders or infectious
PT diseases.

PS Claim 11; SEQ ID NO 918; 413pp; English.

XX
CC The invention describes an isolated polypeptide comprising an amino acid
CC sequence at least 90% identical to: a polypeptide fragment, domain,
CC epitope, or full-length protein of any one of 607 amino acid sequences
CC (I) described in the specification; a polypeptide fragment of (I), or the
CC encoded sequence contained in (II), having biological activity; or a
CC variant, allelic variant, or a species homologue of (I). The polypeptides
CC and nucleic acid molecules are useful for detecting, preventing,
CC diagnosing, prognosticating, treating or ameliorating medical conditions
CC such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
CC Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
CC immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic

Query Match 100.0%; Score 36; DB 8; Length 262;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 17 HPPHG 21

RESULT 37
ABB66594
ID ABB66594 standard; protein; 311 AA.
XX
AC ABB66594;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 26574.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL10697.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 26574; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 311 AA;

Query Match 100.0%; Score 36; DB 4; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 102 HPPHG 106

RESULT 38
AAU51225
ID AAU51225 standard; protein; 324 AA.
XX
AC AAU51225;
XX
DT 27-FEB-2002 (first entry)
XX

DE Propionibacterium acnes immunogenic protein #12121.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59551.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 12420; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 324 AA;

Query Match 100.0%; Score 36; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 254 HPPHG 258

RESULT 39
ABM47744
ID ABM47744 standard; protein; 324 AA.
XX
AC ABM47744;
XX
DT 20-OCT-2003 (first entry)
XX

DE Propionibacterium acnes predicted ORF-encoded polypeptide #12420.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64480.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 12420; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 324 AA;

Query Match 100.0%; Score 36; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||
Db 254 HPPHG 258

RESULT 40
ABU33219
ID ABU33219 standard; protein; 475 AA.
XX

AC ABU33219;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #18746.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Legionella pneumophila.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA37089.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 61143; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 475 AA;

Query Match 100.0%; Score 36; DB 6; Length 475;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
 |||||
Db 412 HPPHG 416

Search completed: June 15, 2005, 14:15:50
Job time : 92 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	231	4	US-09-257-583-9 Sequence 9, Appli
2	36	100.0	262	4	US-09-489-039A-8374 Sequence 8374, Ap
3	36	100.0	480	4	US-09-252-991A-26799 Sequence 26799, A
4	36	100.0	516	4	US-09-252-991A-19788 Sequence 19788, A
5	36	100.0	613	4	US-09-252-991A-22424 Sequence 22424, A
6	36	100.0	1003	4	US-09-252-991A-24069 Sequence 24069, A
7	36	100.0	1447	3	US-09-041-886-25 Sequence 25, Appli
8	36	100.0	1447	5	PCT-US94-05277-2 Sequence 2, Appli
9	30	83.3	7	3	US-08-973-551-19 Sequence 19, Appli
10	30	83.3	16	2	US-08-480-190-29 Sequence 29, Appli
11	30	83.3	16	2	US-08-488-379-29 Sequence 29, Appli
12	30	83.3	16	4	US-08-475-399A-29 Sequence 29, Appli
13	30	83.3	16	4	US-08-077-255A-29 Sequence 29, Appli
14	30	83.3	16	5	PCT-US93-07545-29 Sequence 29, Appli
15	30	83.3	46	4	US-09-270-767-57079 Sequence 57079, A
16	30	83.3	60	4	US-09-248-796A-24991 Sequence 24991, A
17	30	83.3	64	4	US-09-248-796A-26393 Sequence 26393, A
18	30	83.3	69	4	US-09-107-433-2788 Sequence 2788, Ap
19	30	83.3	71	4	US-09-107-532A-5630 Sequence 5630, Ap
20	30	83.3	74	4	US-09-252-991A-24698 Sequence 24698, A
21	30	83.3	98	2	US-08-484-905-59 Sequence 59, Appli
22	30	83.3	98	3	US-08-481-985B-59 Sequence 59, Appli
23	30	83.3	98	3	US-08-370-476-59 Sequence 59, Appli
24	30	83.3	109	4	US-09-513-999C-5250 Sequence 5250, Ap
25	30	83.3	119	1	US-08-469-486-50 Sequence 50, Appli
26	30	83.3	119	2	US-08-469-658-50 Sequence 50, Appli
27	30	83.3	139	4	US-09-621-976-6561 Sequence 6561, Ap

28	30	83.3	142	4	US-09-248-796A-22219 Sequence 22219, A
29	30	83.3	145	4	US-09-252-991A-25194 Sequence 25194, A
30	30	83.3	147	4	US-09-252-991A-16944 Sequence 16944, A
31	30	83.3	147	4	US-09-489-039A-9494 Sequence 9494, Ap
32	30	83.3	162	4	US-09-252-991A-22093 Sequence 22093, A
33	30	83.3	169	4	US-09-461-912A-47 Sequence 47, Appli
34	30	83.3	173	4	US-09-270-767-39233 Sequence 39233, A
35	30	83.3	173	4	US-09-270-767-54450 Sequence 54450, A
36	30	83.3	176	4	US-09-252-991A-24810 Sequence 24810, A
37	30	83.3	180	4	US-09-252-991A-28220 Sequence 28220, A
38	30	83.3	197	4	US-09-252-991A-20521 Sequence 20521, A
39	30	83.3	200	4	US-09-640-211A-743 Sequence 743, App
40	30	83.3	203	4	US-09-270-767-41835 Sequence 41835, A
41	30	83.3	216	1	US-08-451-947-8 Sequence 8, Appli
42	30	83.3	216	2	US-08-424-826A-8 Sequence 8, Appli
43	30	83.3	216	3	US-08-928-694-8 Sequence 8, Appli
44	30	83.3	216	4	US-08-450-842-8 Sequence 8, Appli
45	30	83.3	216	4	US-08-451-390-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-257-583-9
; Sequence 9, Application US/09257583A
; Patent No. 6429362
; GENERAL INFORMATION:
; APPLICANT: Crane, Virginia
; TITLE OF INVENTION: Family Of Maize PR-1 Genes And Promoters
; FILE REFERENCE: 5718-32, 035718/175219
; CURRENT APPLICATION NUMBER: US/09/257,583A
; CURRENT FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays
US-09-257-583-9

Query Match 100.0%; Score 36; DB 4; Length 231;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 174 HPPHG 178

RESULT 2
US-09-489-039A-8374
; Sequence 8374, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8374
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8374

Query Match 100.0%; Score 36; DB 4; Length 262;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
| | | | |
Db 18 HPPHG 22

RESULT 3
US-09-252-991A-26799
; Sequence 26799, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26799
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26799

Query Match 100.0%; Score 36; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
| | | | |
Db 358 HPPHG 362

RESULT 4
US-09-252-991A-19788
; Sequence 19788, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19788
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19788

Query Match 100.0%; Score 36; DB 4; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
| | | | |
Db 63 HPPHG 67

RESULT 5
US-09-252-991A-22424
; Sequence 22424, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22424
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22424

Query Match 100.0%; Score 36; DB 4; Length 613;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
| | | | |
Db 467 HPPHG 471

RESULT 6
US-09-252-991A-24069
; Sequence 24069, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24069
; LENGTH: 1003
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24069

Query Match 100.0%; Score 36; DB 4; Length 1003;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
| | | | |
Db 232 HPPHG 236

RESULT 7
US-09-041-886-25
; Sequence 25, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States

;
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-886-25

Query Match 100.0%; Score 36; DB 3; Length 1447;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 8
PCT-US94-05277-2
; Sequence 2, Application PC/TUS9405277
; GENERAL INFORMATION:
; APPLICANT: Bruskin, Arthur
; APPLICANT: Jarosz, David E.
; APPLICANT: Johnson, Karen
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Vogelstein, Bert
; APPLICANT: Zabrecky, James R.
; TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05277
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42709
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-05277-2

Query Match 100.0%; Score 36; DB 5; Length 1447;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 9
US-08-973-551-19
; Sequence 19, Application US/08973551
; Patent No. 6113902
; GENERAL INFORMATION:
; APPLICANT: Chermann, Jean-Claude
; APPLICANT: Le Contel, Carole
; APPLICANT: Galea, Pascale
; TITLE OF INVENTION: VACCINE AGAINST INFECTIOUS AGENTS HAVING
; TITLE OF INVENTION: AN INTRACELLULAR PHASE, COMPOSITION FOR THE TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF HIV INFECTIONS, ANTIBODIES AND METHOD OF
; TITLE OF INVENTION: DIAGNOSIS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,551
; FILING DATE: 30-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/01006
; FILING DATE: 28-JUN-1996
; APPLICATION NUMBER: FR 9507914
; FILING DATE: 30-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Granados, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 65691/130
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-973-551-19

Query Match 83.3%; Score 30; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 2 HPPH 5

```
RESULT 10
US-08-480-190-29
; Sequence 29, Application US/08480190
; Patent No. 5827516
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,190
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-480-190-29

Query Match      83.3%; Score 30; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      1 HPPH 4

RESULT 11
US-08-488-379-29
; Sequence 29, Application US/08488379
; Patent No. 5880103
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
```

```
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,379
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-488-379-29

Query Match      83.3%; Score 30; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      1 HPPH 4

RESULT 12
US-08-475-399A-29
; Sequence 29, Application US/08475399A
; Patent No. 6509033
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicz, Roman M.
; APPLICANT: Vignali, Dario A.A.
; APPLICANT: Hedley, Mary L.
; APPLICANT: Stern, Lawrence J.
; APPLICANT: Strominger, Jack L.
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 276
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
```


; APPLICATION NUMBER: US/08/475,399A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: 15-JUN-1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00246/168003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-507
; TELEFAX: 617/542-890
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-475-399A-29

Query Match 83.3%; Score 30; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 1 HPPH 4

RESULT 13
US-08-077-255A-29
; Sequence 29, Application US/08077255A
; Patent No. 6696061
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,255A
; FILING DATE: June 15, 1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-077-255A-29
Query Match 83.3%; Score 30; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
|||
Db 1 HPPH 4

RESULT 14
PCT-US93-07545-29
; Sequence 29, Application PC/TUS9307545
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07545
; FILING DATE: 19930811
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-07545-29

Query Match 83.3%; Score 30; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 1 HPPH 4

RESULT 15
US-09-270-767-57079

```
; Sequence 57079, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57079
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57079

Query Match      83.3%; Score 30; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      9 HPPH 12

RESULT 16
US-09-248-796A-24991
; Sequence 24991, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24991
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24991

Query Match      83.3%; Score 30; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      14 HPPH 17

RESULT 17
US-09-248-796A-26393
; Sequence 26393, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26393
```

```
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26393

Query Match      83.3%; Score 30; DB 4; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      47 HPPH 50

RESULT 18
US-09-107-433-2788
; Sequence 2788, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOS
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2788:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...69
; SEQUENCE DESCRIPTION: SEQ ID NO: 2788:
US-09-107-433-2788

Query Match      83.3%; Score 30; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      6 HPPH 9
```

RESULT 19
US-09-107-532A-5630
; Sequence 5630, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5630:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...71
; SEQUENCE DESCRIPTION: SEQ ID NO: 5630:
US-09-107-532A-5630

Query Match 83.3%; Score 30; DB 4; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|
|
|
Db 51 HPPH 54

RESULT 20
US-09-252-991A-24698
; Sequence 24698, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24698
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24698
Query Match 83.3%; Score 30; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HPPH 4
|
|
|
Db 9 HPPH 12
RESULT 21
US-08-484-905-59
; Sequence 59, Application US/08484905
; Patent No. 5976551
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
; TITLE OF INVENTION: Determinant
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS-/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-59
Query Match 83.3%; Score 30; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;

```
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      31 HPPH 34

RESULT 22
US-08-481-985B-59
; Sequence 59, Application US/08481985B
; Patent No. 6011146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Phillipe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,985B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-481-985B-59

Query Match      83.3%; Score 30; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      31 HPPH 34

RESULT 23
US-08-370-476-59
; Sequence 59, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
```

```
; APPLICANT: Kourilsky, Phillipe
; APPLICANT: Lone, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Casrouge, Armanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,476
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,575
; FILING DATE: 07-SEP-1993
; APPLICATION NUMBER: US 08/072,787
; FILING DATE: 06-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-370-476-59

Query Match      83.3%; Score 30; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      31 HPPH 34

RESULT 24
US-09-513-999C-5250
; Sequence 5250, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
```

; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5250
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5250

Query Match 83.3%; Score 30; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 72 HPPH 75

RESULT 25
US-08-469-486-50
; Sequence 50, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-486-50

Query Match 83.3%; Score 30; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 51 HPPH 54

RESULT 26
US-08-469-658-50
; Sequence 50, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-658-50

Query Match 83.3%; Score 30; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 51 HPPH 54

RESULT 27
US-09-621-976-6561
; Sequence 6561, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm

```

; SEQ ID NO 6561
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6561

Query Match      83.3%; Score 30; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      51 HPPH 54

RESULT 28
US-09-248-796A-22219
; Sequence 22219, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22219
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (31)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-22219

Query Match      83.3%; Score 30; DB 4; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      11 HPPH 14

RESULT 29
US-09-252-991A-25194
; Sequence 25194, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25194
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25194

Query Match      83.3%; Score 30; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      55 HPPH 58

RESULT 30
US-09-252-991A-16944
; Sequence 16944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16944
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16944

Query Match      83.3%; Score 30; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      112 HPPH 115

RESULT 31
US-09-489-039A-9494
; Sequence 9494, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9494
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9494

Query Match      83.3%; Score 30; DB 4; Length 147;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      14 YPPHG 18

RESULT 32
US-09-252-991A-22093
; Sequence 22093, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
```



```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22093
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22093

Query Match      83.3%; Score 30; DB 4; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      50 HPPH 53

RESULT 33
US-09-461-912A-47
; Sequence 47, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; TITLE OF INVENTION: differentially expressed genes in disease states
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461,912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,008
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-912A-47

Query Match      83.3%; Score 30; DB 4; Length 169;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      43 HPPH 46

RESULT 34
US-09-270-767-39233
; Sequence 39233, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39233
; LENGTH: 173

```

```

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39233

Query Match      83.3%; Score 30; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      76 HPPH 79

RESULT 35
US-09-270-767-54450
; Sequence 54450, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54450
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54450

Query Match      83.3%; Score 30; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      76 HPPH 79

RESULT 36
US-09-252-991A-24810
; Sequence 24810, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24810
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24810

Query Match      83.3%; Score 30; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      99 HPPH 102

```

```

; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 743
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-743

Query Match      83.3%; Score 30; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      72 HPPH 75

RESULT 40
US-09-270-767-41835
; Sequence 41835, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41835
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41835

Query Match      83.3%; Score 30; DB 4; Length 203;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      9 HPPH 12

Search completed: June 15, 2005, 14:24:20
Job time : 24.5 secs
```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28220
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28220

Query Match      83.3%; Score 30; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      133 HPPH 136

RESULT 38
US-09-252-991A-20521
; Sequence 20521, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20521
; LENGTH: 197
; TYPE: PRT
; ORGANISM: pseudomonas aeruginosa
US-09-252-991A-20521

Query Match      83.3%; Score 30; DB 4; Length 197;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      :||||
Db      4 YPPHG 8

RESULT 39
US-09-640-211A-743
; Sequence 743, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-9
Perfect score: 36
Sequence: 1 HPPHG 5
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	5	14	US-10-074-225A-9
2	36	100.0	20	9	US-09-864-761-41108
3	36	100.0	53	16	US-10-425-115-338760
4	36	100.0	63	9	US-09-764-869-833
5	36	100.0	63	14	US-10-091-504-833
6	36	100.0	63	15	US-10-227-577-833
7	36	100.0	76	15	US-10-424-599-200189
8	36	100.0	92	16	US-10-425-115-345170
9	36	100.0	101	14	US-10-074-225A-6
10	36	100.0	112	15	US-10-424-599-149766
11	36	100.0	117	15	US-10-466-164-37

12	36	100.0	118	14	US-10-023-896-137	Sequence 137, Appl
13	36	100.0	123	16	US-10-437-963-179511	Sequence 179511,
14	36	100.0	126	16	US-10-425-115-261744	Sequence 261744,
15	36	100.0	135	15	US-10-425-114-42907	Sequence 42907, A
16	36	100.0	150	16	US-10-425-115-293118	Sequence 293118,
17	36	100.0	156	16	US-10-425-115-249025	Sequence 249025,
18	36	100.0	168	16	US-10-425-115-265234	Sequence 265234,
19	36	100.0	170	16	US-10-425-115-251226	Sequence 251226,
20	36	100.0	181	15	US-10-424-599-222089	Sequence 222089,
21	36	100.0	230	16	US-10-425-115-317091	Sequence 317091,
22	36	100.0	231	9	US-09-840-479-9	Sequence 9, Appli
23	36	100.0	240	16	US-10-437-963-117204	Sequence 117204,
24	36	100.0	260	15	US-10-425-114-64276	Sequence 64276, A
25	36	100.0	262	11	US-09-764-875-918	Sequence 918, Appl
26	36	100.0	264	15	US-10-425-114-64304	Sequence 64304, A
27	36	100.0	265	16	US-10-425-115-218724	Sequence 218724,
28	36	100.0	313	16	US-10-437-963-155484	Sequence 155484,
29	36	100.0	346	16	US-10-437-963-173538	Sequence 173538,
30	36	100.0	426	15	US-10-425-114-70512	Sequence 70512, A
31	36	100.0	475	15	US-10-282-122A-61143	Sequence 61143, A
32	36	100.0	517	16	US-10-437-963-106725	Sequence 106725,
33	36	100.0	526	14	US-10-074-225A-3	Sequence 3, Appli
34	36	100.0	1498	16	US-10-437-963-200249	Sequence 200249,
35	30	83.3	7	10	US-09-827-345-19	Sequence 39, Appl
36	30	83.3	15	15	US-10-417-895A-39	Sequence 39, Appl
37	30	83.3	25	9	US-09-864-761-40555	Sequence 40555, A
38	30	83.3	25	9	US-09-864-761-47241	Sequence 47241, A
39	30	83.3	38	14	US-10-092-750-46	Sequence 46, Appl
40	30	83.3	38	16	US-10-425-115-263123	Sequence 263123,
41	30	83.3	40	16	US-10-767-701-47944	Sequence 47944, A
42	30	83.3	41	15	US-10-424-599-159525	Sequence 159525,
43	30	83.3	41	15	US-10-424-599-177264	Sequence 177264,
44	30	83.3	43	10	US-09-813-153-129	Sequence 129, Appl
45	30	83.3	43	15	US-10-415-383-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-10-074-225A-9
; Sequence 9, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074, 225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-9

Query Match 100.0%; Score 36; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
|||||
Db 1 HPPHG 5

RESULT 2
US-09-864-761-41108
; Sequence 41108, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41108
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC026236.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.86
US-09-864-761-41108

Query Match 100.0%; Score 36; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5

Db 3 HPPHG 7
RESULT 3
US-10-425-115-338760
; Sequence 338760, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 338760
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_72114C.1.pep
US-10-425-115-338760

Query Match 100.0%; Score 36; DB 16; Length 53;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 27 HPPHG 31

RESULT 4
US-09-764-869-833
; Sequence 833, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 833
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-833

Query Match 100.0%; Score 36; DB 9; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 26 HPPHG 30

```
RESULT 5
US-10-091-504-833
; Sequence 833, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 833
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-504-833
; Query Match 100.0%; Score 36; DB 14; Length 63;
; Best Local Similarity 100.0%; Pred. No. 1.1e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 26 HPPHG 30

RESULT 6
US-10-227-577-833
; Sequence 833, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 833

Qy 1 HPPHG 5
Db 26 HPPHG 30

RESULT 7
US-10-424-599-200189
; Sequence 200189, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200189
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(76)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22798C.1.pap
US-10-424-599-200189
; Query Match 100.0%; Score 36; DB 15; Length 76;
; Best Local Similarity 100.0%; Pred. No. 1.3e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 47 HPPHG 51

RESULT 8
US-10-425-115-345170
; Sequence 345170, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
```

```
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-227-577-833
; Query Match 100.0%; Score 36; DB 15; Length 63;
; Best Local Similarity 100.0%; Pred. No. 1.1e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 26 HPPHG 30

RESULT 7
US-10-424-599-200189
; Sequence 200189, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200189
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(76)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22798C.1.pap
US-10-424-599-200189
; Query Match 100.0%; Score 36; DB 15; Length 76;
; Best Local Similarity 100.0%; Pred. No. 1.3e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 47 HPPHG 51

RESULT 8
US-10-425-115-345170
; Sequence 345170, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
```

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 345170
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77962C.1.pep
US-10-425-115-345170

Query Match      100.0%; Score 36; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      59 HPPHG 63

RESULT 9
US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match      100.0%; Score 36; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      33 HPPHG 37

RESULT 10
US-10-424-599-149766
; Sequence 149766, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149766
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Glycine max
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106260C.1.pep
US-10-424-599-149766

Query Match      100.0%; Score 36; DB 15; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      18 HPPHG 22

RESULT 11
US-10-466-164-37
; Sequence 37, Application US/10466164
; Publication No. US20040058365A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.;
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFFO, Abel;
; APPLICANT: MARWAHA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN,Jr., Edward H.;
; APPLICANT: PERALTA, Careyna H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samantha A.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1215 PCT
; CURRENT APPLICATION NUMBER: US/10/466,164
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/US02/01008
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,065
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,326
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,622
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058365A1 LI:180252.16.orf2:2001JAN12
US-10-466-164-37

Query Match      100.0%; Score 36; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      17 HPPHG 21

RESULT 12
US-10-023-896-137
; Sequence 137, Application US/10023896
```


; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 137
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-896-137

Query Match 100.0%; Score 36; DB 14; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 96 HPPHG 100

RESULT 13
US-10-437-963-179511
; Sequence 179511, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179511
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76967C.1.pep
US-10-437-963-179511

Query Match 100.0%; Score 36; DB 16; Length 123;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 49 HPPHG 53

RESULT 14
US-10-425-115-261744
; Sequence 261744, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261744
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170323C.1.pep
US-10-425-115-261744

Query Match 100.0%; Score 36; DB 16; Length 126;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 85 HPPHG 89

RESULT 15
US-10-425-114-42907
; Sequence 42907, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42907
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700570105_FLI.pep
US-10-425-114-42907

Query Match 100.0%; Score 36; DB 15; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 108 HPPHG 112

RESULT 16
US-10-425-115-293118
; Sequence 293118, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B

```

; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 293118
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(150)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_30403C.1.pep
US-10-425-115-293118

Query Match      100.0%; Score 36; DB 16; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      5 HPPHG 9

RESULT 17
US-10-425-115-249025
; Sequence 249025, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 249025
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_158698C.1.pep
US-10-425-115-249025

Query Match      100.0%; Score 36; DB 16; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      51 HPPHG 55

RESULT 18
US-10-425-115-265234
; Sequence 265234, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 265234
; LENGTH: 168
```

```

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_1734C.1.pep
US-10-425-115-265234

Query Match      100.0%; Score 36; DB 16; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      58 HPPHG 62

RESULT 19
US-10-425-115-251226
; Sequence 251226, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 251226
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(170)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_160703C.1.pep
US-10-425-115-251226

Query Match      100.0%; Score 36; DB 16; Length 170;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
      |||||
Db      149 HPPHG 153

RESULT 20
US-10-424-599-222089
; Sequence 222089, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 222089
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_42576C.1.pep
US-10-424-599-222089
```

Query Match 100.0%; Score 36; DB 15; Length 181;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 |||||
Db 23 HPPHG 27

RESULT 21
US-10-425-115-317091
; Sequence 317091, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 317091
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(230)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_52259C.1.pep
US-10-425-115-317091

Query Match 100.0%; Score 36; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 |||||
Db 170 HPPHG 174

RESULT 22
US-09-840-479-9
; Sequence 9, Application US/09840479
; Patent No. US20010025380A1
; GENERAL INFORMATION:
; APPLICANT: Crane, Virginia
; TITLE OF INVENTION: Family Of Maize PR-1 Genes And Promoters
; FILE REFERENCE: 5718-32, 035718/175219
; CURRENT APPLICATION NUMBER: US/09/840,479
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/257,583
; PRIOR FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays
US-09-840-479-9

Query Match 100.0%; Score 36; DB 9; Length 231;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 |||||
Db 174 HPPHG 178

RESULT 23
US-10-437-963-117204
; Sequence 117204, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117204
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20631C.1.pep
US-10-437-963-117204

Query Match 100.0%; Score 36; DB 16; Length 240;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 |||||
Db 216 HPPHG 220

RESULT 24
US-10-425-114-64276
; Sequence 64276, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64276
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3607-056-B7_FLI.pep
US-10-425-114-64276

Query Match 100.0%; Score 36; DB 15; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 |||||
Db 203 HPPHG 207

RESULT 25

```
US-09-764-875-918
; CURRENT APPLICATION NUMBER: US/09764875
; Sequence 918, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ02
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 918
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-918

Query Match      100.0%; Score 36; DB 11; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
Db      17 HPPHG 21

RESULT 26
US-10-425-114-64304
; Sequence 64304, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64304
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-049-C7_FLI.pep
US-10-425-114-64304

Query Match      100.0%; Score 36; DB 15; Length 264;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
Db      207 HPPHG 211

RESULT 27
US-10-425-115-218724
; Sequence 218724, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 218724
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_131068C.1.pep
US-10-425-115-218724

Query Match      100.0%; Score 36; DB 16; Length 265;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
Db      207 HPPHG 211

RESULT 28
US-10-437-963-155484
; Sequence 155484, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155484
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55242C.1.pep
US-10-437-963-155484

Query Match      100.0%; Score 36; DB 16; Length 313;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HPPHG 5
Db      178 HPPHG 182

RESULT 29
US-10-437-963-173538
; Sequence 173538, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
```

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173538
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71566C.1.pep
US-10-437-963-173538

Query Match 100.0%; Score 36; DB 16; Length 346;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
| | | | |
Db 5 HPPHG 9

RESULT 30
US-10-425-114-70512
; Sequence 70512, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70512
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17050F06_FLI.pep
US-10-425-114-70512

Query Match 100.0%; Score 36; DB 15; Length 426;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
| | | | |
Db 19 HPPHG 23

RESULT 31
US-10-282-122A-61143
; Sequence 61143, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61143
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Legionella pneumophila
US-10-282-122A-61143

Query Match 100.0%; Score 36; DB 15; Length 475;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
| | | | |
Db 412 HPPHG 416

RESULT 32
US-10-437-963-106725
; Sequence 106725, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106725
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11143C.1.pep
US-10-437-963-106725

; Query Match 100.0%; Score 36; DB 16; Length 517;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
| | | | |
Db 382 HPPHG 386

```

; Sequence 19, Application US/09827345
; Publication No. US20030021800A1
; GENERAL INFORMATION:
; APPLICANT: CHERMANN, JEAN-CLAUDE
; APPLICANT: LE CONTEL, CAROLE
; APPLICANT: GALEA, PASCALE
; TITLE OF INVENTION: VACCINE AGAINST INFECTIOUS AGENTS HAVING AN
; TITLE OF INVENTION: INTRACELLULAR PHASE, COMPOSITION FOR THE TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF HIV INFECTIONS, ANTIBODIES AND METHOD OF
; TITLE OF INVENTION: DIAGNOSIS
; FILE REFERENCE: 065691-0216
; CURRENT APPLICATION NUMBER: US/09/827,345
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/599,549
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: PCT/FR96/01006
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 08/973,551
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: FR 95/07914
; PRIOR FILING DATE: 1995-06-30
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-827-345-19

Query Match      83.3%; Score 30; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      2 HPPH 5

RESULT 36
US-10-417-895A-39
; Sequence 39, Application US/10417895A
; Publication No. US20040033569A1
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; APPLICANT: Cappuccilli, Guido
; TITLE OF INVENTION: "DOPING" IN WALK-THROUGH MUTAGENESIS
; FILE REFERENCE: 1551.2002-001
; CURRENT APPLICATION NUMBER: US/10/417,895A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/373,686
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant peptide for second complementarity
; OTHER INFORMATION: determining region of Fv region of an
; OTHER INFORMATION: immunoglobulin
US-10-417-895A-39

Query Match      83.3%; Score 30; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      4 HPPH 7

RESULT 35
US-09-827-345-19
```

```

; Sequence 3, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match      100.0%; Score 36; DB 14; Length 526;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      353 HPPHG 357

RESULT 34
US-10-437-963-200249
; Sequence 200249, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 200249
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_95737C.1.pep
US-10-437-963-200249

Query Match      100.0%; Score 36; DB 16; Length 1498;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      1337 HPPHG 1341

RESULT 35
US-09-827-345-19
```


RESULT 37
US-09-864-761-40555
; Sequence 40555, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40555
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007981.43
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
; OTHER INFORMATION: EST_HUMAN HIT: AI026655.1, EVALUE 3.60e-02
US-09-864-761-40555

Query Match 83.3%; Score 30; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HPPH 4
Db 10 HPPH 13
RESULT 38
US-09-864-761-47241
; Sequence 47241, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47241
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005326.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.47
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.41
; OTHER INFORMATION: EST_HUMAN HIT: H23753.1, EVALUE 8.00e+00
US-09-864-761-47241

Query Match 83.3%; Score 30; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 19 HPPH 22

RESULT 39
US-10-092-750-46
; Sequence 46, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-46

Query Match 83.3%; Score 30; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 34 HPPH 37

RESULT 40
US-10-425-115-263123
; Sequence 263123, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 263123
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171580C.1.pap
US-10-425-115-263123

Query Match 83.3%; Score 30; DB 16; Length 38;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 28 HPPH 31

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-9
Perfect score: 36
Sequence: 1 HPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	118	2 F72674	hypothetical prote
2	36	100.0	176	2 F82478	hypothetical prote
3	36	100.0	235	2 T06168	pathogenesis relat
4	36	100.0	251	2 G72454	probable phosphoe
5	36	100.0	470	2 T46814	gamma-aminobutyrat
6	36	100.0	470	2 B95419	diaminobutyrate-py
7	36	100.0	1427	2 I51669	tumor suppressor -
8	36	100.0	1447	2 A54100	early E3 6.4K prot
9	30	83.3	61	1 ERAD64	hypothetical prote
10	30	83.3	73	2 H71311	hypothetical prote
11	30	83.3	107	2 T47876	hypothetical prote
12	30	83.3	112	2 A03861	hypothetical prote
13	30	83.3	113	2 B72629	hypothetical prote
14	30	83.3	113	2 S66936	probable membrane
15	30	83.3	115	2 T17301	hypothetical prote
16	30	83.3	119	1 MGMSB2	beta-2-microglobul
17	30	83.3	121	2 B72546	hypothetical prote
18	30	83.3	122	2 T46369	hypothetical prote
19	30	83.3	125	2 A05160	hypothetical prote
20	30	83.3	126	2 B69320	conserved hypotHet
21	30	83.3	141	2 F71070	hypothetical prote
22	30	83.3	148	2 S39206	proline-rich prote
23	30	83.3	157	2 A99226	conserved hypotHet
24	30	83.3	159	2 C84299	hypothetical prote
25	30	83.3	167	2 A97760	NADH2 dehydrogenas
26	30	83.3	169	2 JE0149	peripheral benzoda
27	30	83.3	169	2 I38724	mitochondrial benz
28	30	83.3	169	2 I38105	peripheral benzodi
29	30	83.3	169	2 S14257	benzodiazepine rec

30	30	83.3	173	2 T51469	glycine/proline-ri
31	30	83.3	181	2 T52409	uclacyanin II [imp
32	30	83.3	183	2 B97784	formylmethionine d
33	30	83.3	183	2 AC2903	conserved hypotHet
34	30	83.3	203	2 H69937	hypothetical prote
35	30	83.3	207	2 B84315	hypothetical prote
36	30	83.3	208	2 D75004	hypothetical prote
37	30	83.3	209	2 C89005	protein T24A6.3 [i
38	30	83.3	214	2 AF2901	conserved hypotHet
39	30	83.3	214	2 H97676	hypothetical prote
40	30	83.3	216	2 I53100	eHAND - mouse
41	30	83.3	220	2 AH3444	calf thymus ribonu
42	30	83.3	221	2 S25168	CHT28 protein - ch
43	30	83.3	223	2 C87277	conserved hypotHet
44	30	83.3	228	2 T02555	hypothetical prote
45	30	83.3	232	2 G72508	hypothetical prote

ALIGNMENTS

RESULT 1
F72674

hypothetical protein APE0819 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: F72674

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: F72674

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-118 <KAW>

A;Cross-references: UNIPROT:Q9YDU9; DBJ:AP000060; NID:g5104188; PIDN:BAA79798.1; PID:dl
A;Experimental source: strain K1

C;Genetics:

A;Gene: APE0819

C;Superfamily: Aeropyrum pernix hypothetical protein APE0819

Query Match 100.0%; Score 36; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5

|||||

Db 83 HPPHG 87

RESULT 2
F82478

hypothetical protein VCA0284 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: F82478

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: F82478

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-176 <HEI>

A;Cross-references: UNIPROT:Q9KMP0; GB:AE004368; GB:AE003853; NID:g9657677; PIDN:AAF9619;
A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor

C;Genetics:

A;Gene: VCA0284

A;Map position: 2

Query Match		100.0%; Score 36; DB 2; Length 176;	
Best Local Similarity		100.0%; Pred. No. 15;	
Matches		5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HPPHG 5		
Db	51 HPPHG 55		
RESULT 3			
T06168			
pathogenesis related protein - barley			
C;Species: Hordeum vulgare (barley)			
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004			
C;Accession: T06168			
R;Scott, K.J.			
submitted to the EMBL Data Library, June 1994			
A;Reference number: Z15503			
A;Accession: T06168			
A;Status: preliminary; translated from GB/EMBL/DBBJ			
A;Molecule type: mRNA			
A;Residues: 1-235 <SCO>			
A;Cross-references: UNIPROT:P16273; EMBL:X16648; NID:g499073; PIDN:CAA34641.1; PID:g4990			
A;Experimental source: cv. Psakon 4			
Query Match		100.0%; Score 36; DB 2; Length 235;	
Best Local Similarity		100.0%; Pred. No. 20;	
Matches		5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HPPHG 5		
Db	176 HPPHG 180		
RESULT 4			
G72454			
probable phosphoesterase (EC 3.1.-.-) APE2283 [similarity] - Aeropyrum pernix (strain K1			
C;Species: Aeropyrum pernix			
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004			
C;Accession: G72454			
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah			
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K			
DNA Res. 6, 83-101, 1999			
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr			
A;Reference number: A72450; MUID:99310339; PMID:10382966			
A;Accession: G72454			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-251 <KAW>			
A;Cross-references: UNIPROT:Q9Y9K5; DDBJ:AP000064; NID:gs105945; PIDN:BAA81295.1; PID:d1			
A;Experimental source: strain K1			
C;Genetics:			
A;Gene: APE2283			
C;Superfamily: probable phosphoesterase MTH1882; phosphoesterase core homology			
Query Match		100.0%; Score 36; DB 2; Length 251;	
Best Local Similarity		100.0%; Pred. No. 21;	
Matches		5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HPPHG 5		
Db	172 HPPHG 176		
RESULT 5			
T46814			
gamma-aminobutyrate aminotransferase (EC 2.6.1.-) rhbA [imported] - Rhizobium meliloti			
C;Species: Rhizobium meliloti			
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004			
C;Accession: T46814			
R;Lynch, D.; O Connell, M.; O'Brien, J.			
submitted to the EMBL Data Library, December 1998			
A;Description: Cloning and sequence analysis of the Sinorhizobium meliloti 2011 rhizobad			

A;Reference number: Z24097			
A;Accession: T46814			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1-470 <LYN>			
A;Cross-references: UNIPROT:Q9Z3R2; EMBL:AF110737; PIDN:AAD09412.1			
A;Experimental source: strain 2011			
C;Genetics:			
A;Gene: rhbA			
C;Function:			
A;Pathway: siderophore biosynthesis			
C;Superfamily: beta-alanine-pyruvate transaminase			
C;Keywords: aminotransferase			
Query Match		100.0%; Score 36; DB 2; Length 470;	
Best Local Similarity		100.0%; Pred. No. 40;	
Matches		5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HPPHG 5		
Db	406 HPPHG 410		
RESULT 6			
B95419			
diaminobutyrate-pyruvate transaminase (EC 2.6.1.46) RbA [imported] - Sinorhizobium meli			
C;Species: Sinorhizobium meliloti			
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004			
C;Accession: B95419			
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows			
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.			
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001			
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot			
A;Reference number: A95262; MUID:21396509; PMID:11481432			
A;Accession: B95419			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-470 <KUR>			
A;Cross-references: UNIPROT:Q9Z3R2; GB:AE006469; PIDN:AAK65916.1; PID:gl4524429; GSPDB:GN			
A;Experimental source: strain 1021, megaplasmid pSymA			
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,			
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;			
L.; Hyman, R.W.; Jones, T.			
Science 293, 668-672, 2001			
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,			
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.			
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.			
A;Reference number: A96039; MUID:21368234; PMID:11474104			
A;Contents: annotation			
C;Genetics:			
A;Gene: rhbA			
A;Genome: plasmid			
C;Superfamily: beta-alanine-pyruvate transaminase			
C;Keywords: aminotransferase			
Query Match		100.0%; Score 36; DB 2; Length 470;	
Best Local Similarity		100.0%; Pred. No. 40;	
Matches		5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HPPHG 5		
Db	406 HPPHG 410		
RESULT 7			
I51669			
tumor suppressor - African clawed frog			
C;Species: Xenopus laevis (African clawed frog)			
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004			
C;Accession: I51669			
R;Pierceall, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.			
Dev. Biol. 166, 654-665, 1994			
A;Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the			

A;Reference number: I51668; MUID:95113183; PMID:7813784
A;Accession: I51669
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1427 <PIE>
A;Cross-references: UNIPROT:Q91562; EMBL:U10986; NID:g606873; PIDN:AAA70168.1; PID:g6068
C;Genetics:
A;Gene: XDCCA

Query Match 100.0%; Score 36; DB 2; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 1066 HPPHG 1070

RESULT 8
A54100
tumor suppressor protein DCC precursor - human
N;Alternate names: colorectal cancer suppressor DCC
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C;Accession: A54100; A40098
R;Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
Genes Dev. 8, 1174-1183, 1994
A;Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.
A;Reference number: A54100; MUID:95011532; PMID:7926722
A;Accession: A54100
A;Molecule type: mRNA
A;Residues: 1-1447 <HED>
A;Cross-references: UNIPROT:P43146; EMBL:X76132; NID:g453209; PIDN:CAA53735.1; PID:g4532
R;Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilt
Science 247, 49-56, 1990
A;Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.
A;Reference number: A40098; MUID:90100559; PMID:2294591
A;Accession: A40098
A;Molecule type: mRNA
A;Residues: 1-750 <FEA>
A;Cross-references: GB:M32292; NID:g181492; PIDN:AAA35751.1; PID:g181493
C;Genetics:
A;Gene: GDB:DCC
A;Cross-references: GDB:119838; OMIM:120470
A;Map position: 18q21.1-18q21.1
C;Keywords: transmembrane protein; tumor suppressor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 100.0%; Score 36; DB 2; Length 1447;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 9
ERAD64
early E3 6.4K protein - human adenovirus 35
C;Species: Mastadenovirus h35 (human adenovirus 35)
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: B31162
R;Flomenberg, P.R.; Chen, M.; Horwitz, M.S.
J. Virol. 62, 4431-4437, 1988
A;Title: Sequence and genetic organization of adenovirus type 35 early region 3.
A;Reference number: A93039; MUID:89012230; PMID:3172347
A;Accession: B31162
A;Molecule type: DNA
A;Residues: 1-61 <FLO>
A;Cross-references: UNIPROT:P17591; GB:M23195; NID:g516584; PIDN:AAA42436.1; PID:g516586

C;Superfamily: adenovirus early E3 6.4K protein
C;Keywords: early protein

Query Match 83.3%; Score 30; DB 1; Length 61;
Best Local Similarity 80.0%; Pred. No. 60;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 20 YPPHG 24

RESULT 10
H71311
hypothetical protein TP0539 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: H71311
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDor
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: H71311
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-73 <COL>
A;Cross-references: UNIPROT:O83550; GB:AE001229; GB:AE000520; NID:g3322827; PIDN:AAC6553
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0539

Query Match 83.3%; Score 30; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 20 HPPH 23

RESULT 11
T47876
hypothetical protein T4C21.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47876
R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sai
.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24479
A;Accession: T47876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-107 <CHO>
A;Cross-references: UNIPROT:Q9LZZ8; EMBL:AL162295
A;Experimental source: cultivar Columbia; BAC clone T4C21
C;Genetics:
A;Map position: 3
A;Note: T4C21.60

Query Match 83.3%; Score 30; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 102 HPPH 105

RESULT 12
A03861

hypothetical protein F-112 - human adenovirus 2
C;Species: Mastadenovirus h2 (human adenovirus 2)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: G92351; G92352; A03861
R;Gingeras, T.R.; Sciaky, D.; Gelinas, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull
J. Biol. Chem. 257, 13475-13491, 1982
A;Title: Nucleotide sequences from the adenovirus-2 genome.
A;Reference number: A92351; MUID:83056843; PMID:7142161
A;Accession: G92351
A;Molecule type: DNA
A;Residues: 1-112 <GIN>
A;Cross-references: UNIPROT:P03289
R;Alestrom, P.; Akusjarvi, G.; Pettersson, M.; Pettersson, U.
J. Biol. Chem. 257, 13492-13498, 1982
A;Title: DNA sequence analysis of the region encoding the terminal protein and the hypob
A;Reference number: A92352; MUID:83056844; PMID:7142162
A;Accession: G92352
A;Molecule type: DNA
A;Residues: 1-112 <ALE>

Query Match 83.3%; Score 30; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 27 HPPH 30

RESULT 13
B72629
hypothetical protein APE1490 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B72629
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72629
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KAW>
A;Cross-references: UNIPROT:Q9YEW0; DDBJ:AP000061; NID:g5104821; PIDN:BAA80488.1; PID:dl
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1490
C;Superfamily: Aeropyrum pernix hypothetical protein APE1490

Query Match 83.3%; Score 30; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 108 HPPH 111

RESULT 14
S66936
probable membrane protein YOR053w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O2799
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S66936; S66927
R;Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66929
A;Accession: S66936
A;Molecule type: DNA
A;Residues: 1-113 <BOH>
A;Cross-references: UNIPROT:Q08428; EMBL:Z74961; NID:g1420188; PID:e252336; PID:g1420189

A;Experimental source: strain S288C
R;Landt, O.; Hiesel, R.; Unseld, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66907
A;Accession: S66927
A;Molecule type: DNA
A;Residues: 1-8 <LAN>
A;Cross-references: EMBL:Z74961; MIPS:YOR053w
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0005579
A;Map position: 15R
C;Keywords: transmembrane protein
F;3-19/Domain: transmembrane #status predicted <TMM>

Query Match 83.3%; Score 30; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 90 HPPH 93

RESULT 15
T17301
hypothetical protein DKFZp564B147.1 - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17301
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18724
A;Accession: T17301
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-115 <BLO>
A;Cross-references: UNIPROT:Q9UFN1; EMBL:AL117556
A;Experimental source: fetal brain; clone DKFZp564B147
C;Genetics:
A;Note: DKFZp564B147.1
C;Superfamily: human hypothetical protein DKFZp564B147.1

Query Match 83.3%; Score 30; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
Db 55 HPPH 58

RESULT 16
MGMSB2
beta-2-microglobulin precursor - mouse
N;Alternate names: class I histocompatibility antigen beta chain
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1981 #sequence_revision 03-Aug-1984 #text change 09-Jul-2004
C;Accession: A90822; A93889; B93889; A12344; I79322; I58971; I63179; I63180; I63178; I633
R;Parnes, J.R.; Seidman, J.G.
Cell 29, 661-669, 1982
A;Title: Structure of wild-type and mutant mouse beta-2-microglobulin genes.
A;Reference number: A90822; MUID:83001942; PMID:6180834
A;Accession: A90822
A;Molecule type: DNA
A;Residues: 1-119 <PRN>
A;Cross-references: UNIPROT:P01887
A;Experimental source: strain BALB/c
R;Gates III, F.T.; Coligan, J.E.; Kindt, T.J.
Proc. Natl. Acad. Sci. U.S.A. 78, 554-558, 1981
A;Title: Complete amino acid sequence of murine beta-2-microglobulin: structural evidence
A;Reference number: A93889; MUID:81223779; PMID:6165997
A;Accession: A93889

A;Molecule type: protein
A;Residues: 21-119 <GAT>
A;Experimental source: strain BALB/c
A;Accession: B93889
A;Molecule type: protein
A;Residues: 21-104,'A',106-119 <GA2>
A;Experimental source: strain C57BL/6
R;Appella, E.; Tanigaki, N.; Natori, T.; Pressman, D.
Biochem. Biophys. Res. Commun. 70, 425-430, 1976
A;Title: Partial amino acid sequence of mouse beta-2-microglobulin.
A;Reference number: A12344; MUID:76231543; PMID:59595
A;Accession: A12344
A;Molecule type: protein
A;Residues: 21-48,'E',50-53,'X',55,'Z',57,'BL',60 <APP>
A;Experimental source: strain A/J
R;Parnes, J.R.; Velan, B.; Felsenfeld, A.; Ramanathan, L.; Ferrini, U.; Appella, E.; Seidman
Proc. Natl. Acad. Sci. U.S.A. 78, 2253-2257, 1981
A;Title: Mouse beta-2-microglobulin cDNA clones: A screening procedure for cDNA clones of
A;Reference number: I58971; MUID:81223857; PMID:6166005
A;Accession: I79322
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 86-104,'A',106-119 <PAR>
A;Cross-references: GB:M10416; NID:G199147; PIDN:AAA39537.1; PID:G387432
A;Accession: I58971
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 31-57 <RE2>
A;Cross-references: GB:M10389; NID:G199146; PIDN:AAA39536.1; PID:G199149
R;Hermel, E.; Robinson, P.J.; Lindahl, K.F.
Immunogenetics 38, 106-116, 1993
A;Title: Sequence divergence of B2m alleles of wild Mus musculus and Mus spretus implies
A;Reference number: I49537; MUID:93246292; PMID:8482575
A;Accession: I63179
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-7,'I',9-63,'E',65-119 <PA1>
A;Cross-references: GB:M84364; NID:G199575; PIDN:AAA39668.1; PID:G199576
A;Accession: I63180
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-53,'Q',55-104,'V',106-119 <PA2>
A;Cross-references: GB:M84365; NID:G199577; PIDN:AAA39669.1; PID:G199578
A;Accession: I63178
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 24-100,'T',102-104,'V',106-115 <PA3>
A;Cross-references: GB:M84363; NID:G199573; PIDN:AAA39667.1; PID:G199574
A;Accession: I63181
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 24-53,'Q',55-104,'V',106-115 <PA4>
A;Cross-references: GB:M84366; NID:G199579; PIDN:AAA39670.1; PID:G199580
A;Accession: I63182
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-104,'V',106-119 <PA5>
A;Cross-references: GB:M84367; NID:G199581; PIDN:AAA39671.1; PID:G199582
R;Daniel, F.; Morello, D.; Le Bail, O.; Chambon, P.; Cayre, Y.; Kourilsky, P.
EMBO J. 2, 1061-1065, 1983
A;Title: Structure and expression of the mouse beta 2-microglobulin gene isolated from
A;Reference number: I48262; MUID:84028577; PMID:6354707
A;Accession: I48262
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-119 <RES>
A;Cross-references: EMBL:X01838; NID:G50104; PIDN:CAA25970.1; PID:G50105
R;Seubert, P.; Vigo-Pelfrey, C.; Esch, F.; Lee, M.; Dovey, H.; Davis, D.; Sinha, S.; Schenk
Nature 359, 325-327, 1992
A;Title: Isolation and quantification of soluble Alzheimer's beta-peptide from biological
A;Reference number: S29076; MUID:93024877; PMID:1406936
A;Accession: S29077

A;Molecule type: protein
A;Residues: 'X',22-30,'X',32-44 <SEU>
A;Note: this protein was found as a minor contaminant during the purification of proteins
C;Genetics:
A;Introns: 23/1; 116/1
C;Complex: heterodimer with MHC class I histocompatibility antigen alpha chain
C;Function:
A;Description: necessary for the expression of MHC class I histocompatibility antigen
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: heterodimer
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-119/Product: beta-2-microglobulin #status experimental <MAT>
F;38-102/Domain: immunoglobulin homology <IMM>
F;45-100/Disulfide bonds: #status predicted
Query Match 83.3%; Score 30; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HPPH 4
Db 51 HPPH 54
RESULT 17
B72546
hypothetical protein APE1654 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B72546
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara,
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72546
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-121 <KAW>
A;Cross-references: UNIPROT:Q9YBE3; DDBJ:AP000062; NID:G5105244; PIDN:BAA80655.1; PID:G51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1654
C;Superfamily: Aeropyrum pernix hypothetical protein APE1654
Query Match 83.3%; Score 30; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HPPH 4
Db 36 HPPH 39
RESULT 18
T46369
hypothetical protein DKFZp434I0118.1 - human
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46369
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46369
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-122 <AAA>
A;Cross-references: UNIPROT:Q9NT54; EMBL:AL137525
A;Experimental source: adult testis; clone DKFZp434I0118
C;Genetics:
A;Note: DKFZp434I0118.1
C;Superfamily: human hypothetical protein DKFZp434I0118.1

Query Match				83.3%; Score 30; DB 2; Length 122;			
Best Local Similarity				100.0%; Pred. No. 1.2e+02;			
Matches				4; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1 HPPH 4						
Db	14 HPPH 17						
RESULT 19							
A05160							
hypothetical protein 49 - fruit fly (Drosophila melanogaster)							
C;Species: Drosophila melanogaster							
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 16-Feb-1997							
C;Accession: A05160							
R;O'Connell, P.; Rosbash, M.							
Nucleic Acids Res. 12, 5495-5513, 1984							
A;Title: Sequence, structure, and codon preference of the Drosophila ribosomal protein 4							
A;Reference number: A93527; MUID:84272233; PMID:6087289							
A;Accession: A05160							
A;Molecule type: DNA							
A;Residues: 1-125 <OCON>							
C;Genetics:							
A;Gene: FlyBase:M(3)99D							
A;Cross-references: FlyBase:FBgn0002626							
A;Map position: 3R, 99D							
Query Match				83.3%; Score 30; DB 2; Length 125;			
Best Local Similarity				100.0%; Pred. No. 1.2e+02;			
Matches				4; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1 HPPH 4						
Db	32 HPPH 35						
RESULT 20							
B69320							
conserved hypothetical protein AF0562 - Archaeoglobus fulgidus							
C;Species: Archaeoglobus fulgidus							
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004							
C;Accession: B69320							
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson							
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.							
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.							
Nature 390, 364-370, 1997							
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.							
Smith, H.O.; Woese, C.R.; Venter, J.C.							
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo							
A;Reference number: A69250; MUID:98049343; PMID:9389475							
A;Accession: B69320							
A;Status: preliminary; nucleic acid sequence not shown; translation not shown							
A;Molecule type: DNA							
A;Residues: 1-126 <KLE>							
A;Cross-references: UNIPROT:O29691; GB:AE001066; GB:AE000782; NID:g2689389; PIDN:AAB9067							
Query Match				83.3%; Score 30; DB 2; Length 126;			
Best Local Similarity				100.0%; Pred. No. 1.2e+02;			
Matches				4; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1 HPPH 4						
Db	82 HPPH 85						
RESULT 21							
F71070							
hypothetical protein PH1258 - Pyrococcus horikoshii							
C;Species: Pyrococcus horikoshii							
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004							
C;Accession: F71070							
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin							
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi							

DNA Res. 5, 55-76, 1998							
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon							
A;Reference number: A71000; MUID:98344137; PMID:9679194							
A;Accession: F71070							
A;Status: preliminary; nucleic acid sequence not shown; translation not shown							
A;Molecule type: DNA							
A;Residues: 1-141 <KAW>							
A;Cross-references: UNIPROT:O58995; GB:AP000005; NID:g3236132; PIDN:BAA30360.1; PID:d1033							
A;Experimental source: strain OT3							
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank							
C;Genetics:							
A;Gene: PHI258							
Query Match				83.3%; Score 30; DB 2; Length 141;			
Best Local Similarity				100.0%; Pred. No. 1.4e+02;			
Matches				4; Conservative	0; Mismatches	0; Indels	0; Gaps
QY	1 HPPH 4						
Db	65 HPPH 68						
RESULT 22							
S39206							
proline-rich protein V-beta 1 precursor - rat							
C;Species: Rattus norvegicus (Norway rat)							
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004							
C;Accession: A53118; S39206							
R;Court, Y.; Rosinski-Chupin, I.; Rougeon, F.							
J. Biol. Chem. 269, 520-527, 1994							
A;Title: A new proline-rich protein precursor expressed in the salivary glands of the rat							
A;Reference number: A53118; MUID:94103265; PMID:8276845							
A;Accession: A53118							
A;Status: translated from GB/EMBL/DBJ							
A;Molecule type: mRNA							
A;Residues: 1-148 <COU>							
A;Cross-references: UNIPROT:Q64371; EMBL:X74229; NID:g433616; PIDN:CAA52300.1; PID:g43361							
A;Note: submitted to the EMBL Data Library, July 1993							
C;Genetics:							
A;Gene: VCS-beta1							
A;Introns: 18/3							
C;Superfamily: proline-rich peptide P-B							
C;Keywords: glycoprotein							
F;1-18/Domain: signal sequence #status predicted <SIG>							
F;19-148/Product: proline-rich protein V-beta 1 #status predicted <MAT>							
F;133,143/Binding site: carbohydrate (Asn) (covalent) #status predicted							
Query Match				83.3%; Score 30; DB 2; Length 148;			
Best Local Similarity				100.0%; Pred. No. 1.5e+02;			
Matches				4; Conservative	0; Mismatches	0; Indels	0; Gaps
QY	1 HPPH 4						
Db	97 HPPH 100						
RESULT 23							
A99226							
conserved hypothetical protein [imported] - Sulfolobus solfataricus							
C;Species: Sulfolobus solfataricus							
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004							
C;Accession: A99226							
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v							
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.							
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.							
submitted to GenBank, April 2001							
A;Description: Sulfolobus solfataricus complete genome.							
A;Reference number: A99139							
A;Accession: A99226							
A;Status: preliminary							
A;Molecule type: DNA							
A;Residues: 1-157 <KUR>							
A;Cross-references: UNIPROT:Q9UXF4; GB:AE006641; NID:g13813940; PIDN:AAK41064.1; GSPDB:G							

C;Genetics:
A;Gene: SSO0767
C;Superfamily: conserved hypothetical protein ybhB

Query Match 83.3%; Score 30; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 109 HPPH 112

RESULT 24
C84299
hypothetical protein Vng1453h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84299
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonka, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84299
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-159 <STO>
A;Cross-references: UNIPROT:Q9HPV8; GB:AE004437; NID:g10580949; PIDN:AAG19759.1; GSPDB:G
C;Genetics:
A;Gene: VNG1453H

Query Match 83.3%; Score 30; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 98 HPPH 101

RESULT 25
A97760
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97760
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: A97760
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-167 <KUR>
A;Cross-references: UNIPROT:Q92ID9; GB:AE006914; PIDN:AAL03019.1; PID:g15619555; GSPDB:G
C;Genetics:
A;Gene: nuoE
C;Superfamily: NADH dehydrogenase (ubiquinone) I chain E; NADH dehydrogenase (ubiquinone
C;Keywords: oxidoreductase; NAD

Query Match 83.3%; Score 30; DB 2; Length 167;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 25 YPPHG 29

RESULT 26

JE0149
peripheral benzodiazepine receptor - human
C;Species: Homo sapiens (man)
C;Date: 10-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 12-Jul-2004
C;Accession: JE0149
R;Xu, G.H.; Hu, B.R.; Luo, C.Q.; Chen, R.Z.; Qiu, P.X.
Chinese Biochem. J. 13, 423-427, 1997
A;Title: Cloning of the cDNAs encoding human peripheral benzodiazepine receptor and its n
A;Reference number: JE0149
A;Accession: JE0149
A;Molecule type: mRNA
A;Residues: 1-169 <XUA>
C;Superfamily: peripheral-type benzodiazepine receptor/signal transduction protein Tspo

Query Match 83.3%; Score 30; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 43 HPPH 46

RESULT 27
I38724
mitochondrial benzodiazepine receptor - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 12-Jul-2004
C;Accession: I38724
R;Yakovlev, A.G.; Ruffo, M.; Jurka, J.; Krueger, K.E.
Gene 155, 201-205, 1995
A;Title: Comparison of repetitive elements in the third intron of human and rodent mitoch
A;Reference number: I38724; MUID:95237610; PMID:7721091
A;Accession: I38724
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <RES>
A;Cross-references: UNIPROT:P30536; EMBL:U12421; NID:g529945; PIDN:AAA83252.1; PID:g52994
C;Genetics:
A;Gene: MBR
A;Introns: 61/2; 107/3
C;Superfamily: peripheral-type benzodiazepine receptor/signal transduction protein Tspo
C;Keywords: mitochondrion

Query Match 83.3%; Score 30; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 43 HPPH 46

RESULT 28
I38105
peripheral benzodiazepine receptor - human
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 12-Jul-2004
C;Accession: I38105
R;Lin, D.; Chang, Y.J.; Strauss, J.F.; Miller, W.L.
Genomics 18, 643-650, 1993
A;Title: The human peripheral benzodiazepine receptor gene: cloning and characterization
A;Reference number: A49361; MUID:94140364; PMID:8307574
A;Accession: I38105
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <RES>
A;Cross-references: UNIPROT:Q13850; GB:I21954; NID:g483405; PIDN:AAA18228.1; PID:g488425
C;Genetics:
A;Introns: 61/2; 107/3
C;Superfamily: peripheral-type benzodiazepine receptor/signal transduction protein Tspo

Query Match 83.3%; Score 30; DB 2; Length 169;

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;
QY 1 HPPH 4
Db 43 HPPH 46

RESULT 29
S14257
benzodiazepine receptor, peripheral-type - human
C;Species: Homo sapiens (man)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 12-Jul-2004
C;Accession: S14257
R;Riond, J.; Mattei, M.G.; Kaghad, M.; Dumont, X.; Guillemot, J.C.; le Fur, G.; Caput, D
Eur. J. Biochem. 195, 305-311, 1991
A;Title: Molecular cloning and chromosomal localization of a human peripheral-type benzodiazepine receptor
A;Reference number: S14257; MUID:91146565; PMID:1847678
A;Accession: S14257
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-169 <RIO>
A;Cross-references: UNIPROT:P30536; GB:M36035; NID:g184333; PIDN:AAA03652.1; PID:g306883
C;Genetics:
A;Gene: GDB:BZRP
A;Cross-references: GDB:127347; OMIM:109610
A;Map position: 22q13.31-22qter
C;Superfamily: peripheral-type benzodiazepine receptor/signal transduction protein TsPO
C;Keywords: mitochondrion; transmembrane protein

Query Match 83.3%; Score 30; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
Db 43 HPPH 46

RESULT 30
T51469
glycine/proline-rich protein - Arabidopsis thaliana
N;Alternate names: protein K10A8_130
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51469
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51469
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <SAT>
A;Cross-references: UNIPROT:Q9LFS9; EMBL:AL391151
A;Experimental source: cultivar Columbia; BAC clone K10A8
C;Genetics:
A;Map position: 5
A;Introns: 97/1
A;Note: K10A8_130

Query Match 83.3%; Score 30; DB 2; Length 173;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 25 YPPHG 29

RESULT 31
T52409
ucLacyanin II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52409
R;Nersissian, A.M.; Mehrabian, Z.B.; Nalbandyan, R.M.; Hart, P.J.; Fraczekiewicz, G.; Cze
Protein Sci. 5, 2184-2192, 1996
A;Title: Cloning, expression, and spectroscopic characterization of Cucumis sativus stell
A;Reference number: Z17046; MUID:97084803; PMID:8931137
A;Accession: T52409
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-181 <NER>
A;Cross-references: UNIPROT:O80517; EMBL:U76299; PIDN:AAC32039.1
C;Superfamily: stellacyanin

Query Match 83.3%; Score 30; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
Db 156 HPPH 159

RESULT 32
B97784
formylmethionine deformylase (EC.3.5.1.31) - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97784
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: B97784
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-183 <KUR>
A;Cross-references: UNIPROT:Q92HU7; GB:AE006914; PIDN:AAL03212.1; PID:g15619763; GSPDB:GN
C;Genetics:
A;Gene: def2
C;Superfamily: peptide deformylase
C;Keywords: hydrolase

Query Match 83.3%; Score 30; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
Db 172 HPPH 175

RESULT 33
AC2903
conserved hypothetical protein Atu2660 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AC2903
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, F
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2903
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-183 <KUR>
A;Cross-references: UNIPROT:Q8UC38; GB:AE008688; PIDN:AAL43641.1; PID:g17741163; GSPDB:GN
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2660

A;Map position: circular chromosome

Query Match 83.3%; Score 30; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 43 HPPH 46

RESULT 34
H69937
hypothetical protein ypjP - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69937
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, A.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallenbach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seronakeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: H69937
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-203 <KUN>
A;Cross-references: UNIPROT:P54172; GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB14102.
A;Experimental source: strain 168
C;Genetics:
A;Gene: ypjP
C;Superfamily: Bacillus subtilis hypothetical protein ypjP

Query Match 83.3%; Score 30; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 162 HPPH 165

RESULT 35
B84315
hypothetical protein Vng1625h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84315
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84315
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-207 <STO>
A;Cross-references: UNIPROT:Q9HP11; GB:AE004437; NID:g10581098; PIDN:AAG19886.1; GSPDB:C A;Gene: VNG1625H

Query Match 83.3%; Score 30; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 23 HPPH 26

RESULT 36
D75004
hypothetical protein PAB1037 - Pyrococcus abyssi (strain Orsay)
N;Alternate names: yafJ homolog
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: D75004
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc A;Reference number: A75001
A;Accession: D75004
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-208 <KAW>
A;Cross-references: UNIPROT:Q9UYD8; GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50474 A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1037

Query Match 83.3%; Score 30; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 198 HPPH 201

RESULT 37
C89005
protein T24A6.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: C89005
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and A;Accession: C89005
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-209 <STO>
A;Cross-references: GB:chr_V; PIDN:AAC17789.1; PID:g3168937; GSPDB:GN00023; CESP:T24A6.3 C;Genetics:
A;Gene: T24A6.3
A;Map position: 5

Query Match 83.3%; Score 30; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 148 HPPH 151

RESULT 38
AF2901
conserved hypothetical protein Atu2647 [imported] - Agrobacterium tumefaciens (strain C58 C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AF2901
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-214 <KUR>
A;Cross-references: UNIPROT:Q8UC51; GB:AE008688; PIDN:AAL43628.1; PID:gl7741149; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2647
A;Map position: circular chromosome

Query Match 83.3%; Score 30; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 17 HPPH 20

RESULT 39
H97676
hypothetical protein AGR_C_4799 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97676
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H97676
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-214 <KUR>
A;Cross-references: UNIPROT:Q8UC51; GB:AE007869; PIDN:AAK88369.1; PID:gl51517854; GSPDB:G
C;Genetics:
A;Gene: AGR_C_4799
A;Map position: circular chromosome

Query Match 83.3%; Score 30; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 17 HPPH 20

RESULT 40
I53100
eHAND - mouse
C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I53100
R;Cserjesi, P.; Brown, D.; Lyons, G.E.; Olson, E.N.
Dev. Biol. 170, 664-678, 1995
A;Title: Expression of the novel basic helix-loop-helix gene eHAND in neural crest deriv
A;Reference number: I53100; MUID:95377552; PMID:7649392
A;Accession: I53100
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-216 <RES>
A;Cross-references: GB:S79216; NID:gl086931; PIDN:AAB35104.1; PID:gl086932

Query Match 83.3%; Score 30; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
Db 17 HPPH 20

Search completed: June 15, 2005, 14:22:49
Job time : 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-9
Perfect score: 36
Sequence: 1 HPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	36	100.0	62	2	Q15952	Q15952 homo sapien
2	36	100.0	108	2	Q655P3	Q655P3 oryza sativ
3	36	100.0	118	2	Q9YDU9	Q9ydu9 aeropyrum p
4	36	100.0	171	2	Q7PG70	Q7pg70 anopheles g
5	36	100.0	176	2	Q9KMP0	Q9kmp0 vibrio chol
6	36	100.0	193	2	Q8FJG7	Q8fig7 escherichia
7	36	100.0	195	2	Q52719	Q52719 rhodobacter
8	36	100.0	214	2	Q8PQU3	Q8pqu3 xanthomonas
9	36	100.0	234	2	Q7USZ6	Q7usz6 rhodopirell
10	36	100.0	235	1	PRPX HORVU	P16273 hordeum vul
11	36	100.0	240	2	Q9XHX3	Q9xhx3 oryza sativ
12	36	100.0	240	2	Q9XHX6	Q9xhx6 oryza sativ
13	36	100.0	251	2	Q9Y9K5	Q9y9k5 aeropyrum p
14	36	100.0	277	2	P96427	P96427 rhodococcus
15	36	100.0	311	2	Q9YJN5	Q9yjn5 equid herpe
16	36	100.0	321	2	Q8FT73	Q8ft73 corynebacte
17	36	100.0	355	2	Q73RX1	Q73rx1 mycobacteri
18	36	100.0	399	2	Q6MYV5	Q6myv5 aspergillus
19	36	100.0	449	2	Q896I5	Q896i5 clostridium
20	36	100.0	470	1	RHBA RHIME	Q9z3r2 r diaminobu
21	36	100.0	491	2	Q73TJ7	Q73tj7 mycobacteri
22	36	100.0	493	2	Q6P5Y4	Q6p5y4 homo sapien
23	36	100.0	503	2	Q9W259	Q9w259 drosophila
24	36	100.0	526	1	HRG RABIT	Q28640 oryctolagus
25	36	100.0	560	2	Q9YHE8	Q9yhe8 brachydanio
26	36	100.0	581	2	Q87RZ6	Q87rz6 vibrio para
27	36	100.0	583	2	Q8D843	Q8d843 vibrio vuln
28	36	100.0	587	2	Q7MMD1	Q7mmd1 vibrio vuln
29	36	100.0	601	2	Q75HV8	Q75hv8 oryza sativ
30	36	100.0	723	2	Q86BB8	Q86bb8 drosophila
31	36	100.0	736	2	Q6C276	Q6c276 yarrowia li

32	36	100.0	773	2	Q95R83	Q95r83 drosophila
33	36	100.0	938	2	Q86BB9	Q86bb9 drosophila
34	36	100.0	1075	2	Q6XLI4	Q6xli4 ciona intes
35	36	100.0	1427	2	Q91562	Q91562 xenopus lae
36	36	100.0	1445	2	Q63155	Q63155 rattus norv
37	36	100.0	1447	1	DCC_HUMAN	P43146 homo sapien
38	36	100.0	1447	1	DCC_MOUSE	P70211 mus musculu
39	36	100.0	2013	2	Q6SLE5	Q6sle5 cochliobolu
40	30	83.3	16	2	Q9S8A0	Q9s8a0 pinus monti
41	30	83.3	30	2	Q9V2V2	Q9v2v2 methanobact
42	30	83.3	30	2	Q7LYR7	Q7lyr7 plasmid pme
43	30	83.3	54	2	Q8KM82	Q8km82 mycoplasma
44	30	83.3	55	2	Q9NDZ7	Q9ndz7 leishmania
45	30	83.3	61	1	E306_ADB35	P17591 human adeno

ALIGNMENTS

RESULT 1
Q15952
ID Q15952 PRELIMINARY; PRT; 62 AA.
AC Q15952;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Agrin (Fragment).
GN Name=AGRN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92407628; PubMed=1326608;
RA Rupp F., Ozcelik T., Linial M., Peterson K., Francke U., Scheller R.;
RT "Structure and chromosomal localization of the mammalian agrin gene.";
RL J. Neurosci. 12:3535-3544(1992).
DR EMBL; S44195; AAB23327.1; --.
DR HSSP; P21674; 1LR7.
DR GO; GO:0005605; C:basal lamina; TAS.
DR GO; GO:0007268; P:synaptic transmission; NAS.
DR InterPro; IPR003645; FOL_N.
DR InterPro; IPR002350; Prot_inh_Kazal.
DR Pfam; PF00050; Kazal_1; 1.
DR SMART; SM00274; FOLN_1.
DR SMART; SM00280; KAZAL; 1.
FT NON TER 62
SQ SEQUENCE 62 AA; 6344 MW; 4CD1653E6B8F09E8 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||
Db 26 HPPHG 30

RESULT 2
Q655P3
ID Q655P3 PRELIMINARY; PRT; 108 AA.
AC Q655P3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B1011A07.27.
GN Name=B1011A07.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;

```
RN
RP SEQUENCE FROM N.A.
RX PubMed=12447438;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003722; BAD45462.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11886 MW; 6B93CD836E85C0A7 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
29 HPPHG 33

RESULT 3
Q9YDU9
ID Q9YDU9 PRELIMINARY; PRT; 118 AA.
AC Q9YDU9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE0819.
GN OrderedLocusNames=APE0819;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79798.1; -.
DR PIR; F72674; F72674.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 118 AA; 12953 MW; 6134441D6406CDB0 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
83 HPPHG 87

RESULT 4
```

```
Q7PG70
ID Q7PG70 PRELIMINARY; PRT; 171 AA.
AC Q7PG70;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000022414 (Fragment).
GN Name=ENSANGG00000020585;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100847; EAA45048.1; -.
FT NON_TER 1 1
FT NON_TER 171 171
SQ SEQUENCE 171 AA; 18450 MW; 1A135D70C5F14B67 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db |||||
111 HPPHG 115

RESULT 5
Q9KMP0
ID Q9KMP0 PRELIMINARY; PRT; 176 AA.
AC Q9KMP0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein VCA0284.
GN OrderedLocusNames=VCA0284;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004368; AAF96193.1; -.
DR PIR; F82478; F82478.
DR TIGR; VCA0284; -.
DR InterPro; IPR008727; PAAR.
DR Pfam; PF05488; PAAR_motif; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 176 AA; 18431 MW; 79A633E68C447515 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Qy 1 HPPHG 5
Db 199 HPPHG 203

RESULT 9
Q7USZ6 PRELIMINARY; PRT; 234 AA.
AC Q7USZ6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RB4213;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294140; CAD73646.1; --
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 234 AA; 26924 MW; 03A974166BAA3FEC CRC64;

Query Match 100.0%; Score 36; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 135 HPPHG 139

RESULT 10
PRPX HORVU STANDARD; PRT; 235 AA.
AC P16273;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pathogen-related protein.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Psakon 4; TISSUE=Leaf;
RX MEDLINE=90067951; PubMed=2587271;
RA Jutidamrongphan W., Mackinnon G., Manners J.M., Scott K.J.;
RT "Sequence of a near-full length cDNA clone for a mRNA of barley
induced by fungal infection.";
RL Nucleic Acids Res. 17:9478-9478 (1989).
RN [2]
RP REVISIONS.
RC STRAIN=cv. Psakon 4; TISSUE=Leaf;
RA Scott K.J.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
CC -!- INDUCTION: By fungus Bipolaris sorokiniana infection.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16648; CAA34641.1; --
DR PIR; S09604; S09604.
DR PIR; T06168; T06168.
KW Pathogenesis-related protein; Plant defense.
SQ SEQUENCE 235 AA; 26293 MW; 995136DD24A9CD3C CRC64;

Query Match 100.0%; Score 36; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 176 HPPHG 180

RESULT 11
Q9XHX3 PRELIMINARY; PRT; 240 AA.
AC Q9XHX3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative pathogenesis related protein.
GN Name=OSJNBa0049B20.20; Synonyms=P0034C11.23;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R., Benito M.-I., Lin X., Mason T.M., Umayam L., Shea T.P.,
RA Fujii C.Y., Shen M., Fraser C.M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316 (2002).
DR EMBL; AC007789; AAD38293.1; --
DR EMBL; AP002865; BAB18332.1; --
DR Gramene; Q9XHX3; --
SQ SEQUENCE 240 AA; 27196 MW; 4ABC7B300FAFEBD6 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 177 HPPHG 181

RESULT 12
Q9XHX6

```
ID Q9XHX6 PRELIMINARY; PRT; 240 AA.
AC Q9XHX6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative pathogenesis related protein.
GN Name=OSJNBa0049B20.17; SynonymS=P0034C11.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R., Benito M.-I., Lin X., Mason T.M., Umayam L., Shea T.P.,
RA Fujii C.Y., Shen M., Fraser C.M.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AC007789; AAD38290.1; -.
DR EMBL; AP002865; BAB18330.1; -.
DR Gramene; Q9XHX6; -.
SQ SEQUENCE 240 AA; 27146 MW; 86C7D7A6F409AD7E CRC64;

Query Match 100.0%; Score 36; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 178 HPPHG 182

RESULT 13
Q9Y9K5 PRELIMINARY; PRT; 251 AA.
ID Q9Y9K5
AC Q9Y9K5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE2283.
GN OrderedLocusNames=APE2283;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
```

```
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000064; BAA81295.1; -.
DR PIR; G72454; G72454.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR004843; M-pesterase.
DR Pfam; PF00149; Metallophos; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 251 AA; 27075 MW; 9C30FE5292F51B60 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 172 HPPHG 176

RESULT 14
P96427 PRELIMINARY; PRT; 277 AA.
ID P96427
AC P96427;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AraC-type regulatory protein.
GN Name=fasR;
OS Rhodococcus fascians.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D188;
RX MEDLINE=20461232; PubMed=11004184;
RX DOI=10.1128/JB.182.20.5832-5840.2000;
RA Temmerman W., Vereecke D., Dreesen R., Van Montagu M., Holsters M.,
RA Goethals K.;
RT "Leafy gall formation is controlled by fasR, an AraC-type regulatory
RT gene in Rhodococcus fascians.";
RL J. Bacteriol. 182:5832-5840(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=D188;
RA Vereecke D.M.;
RL EMBL; Y09820; CAA70950.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF00165; HTH_AraC; 2.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 277 AA; 30866 MW; 071DB9DEFC11E113 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 155 HPPHG 159

RESULT 15
Q9YJN5 PRELIMINARY; PRT; 311 AA.
ID Q9YJN5
```


RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14998527; DOI=10.1016/j.fgb.2003.12.003;
RA Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,
RA Fosker N., Fraser A., Harris D., Larke N., Murphy L., Humphray S.,
RA O'Neil S., Pertea M., Price C., Rabinowitsch E., Rajandream M-A.,
RA Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,
RA Denning D.W., Barrell B., Hall N.,
RT "Insight into the genome of Aspergillus fumigatus: analysis of a 922
RT kb region encompassing the nitrate assimilation gene cluster.";
RL Fungal Genet. Biol. 41:443-453(2004).
CC -|- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; BX649605; CAE47904.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
SQ SEQUENCE 399 AA; 42514 MW; D23C914E172F615C CRC64;

Query Match 100.0%; Score 36; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 310 HPPHG 314

RESULT 19
Q896I5
ID Q896I5 PRELIMINARY; PRT; 449 AA.
AC Q896I5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Rnfc/nqrf.
GN Name=rnfc; OrderedLocusNames=CTC01019;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;
RA Brueggemann H., Baeumer S., Fricke W.F., Wierer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015939; AA035605.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR011538; Complex1_51K_dom.
DR InterPro; IPR009051; Helical_Ferredxn.
DR InterPro; IPR010208; Rnfc.
DR InterPro; IPR011054; Rudmnt_hyb_motif.
DR Pfam; PF01512; Complex1_51k; 1.
DR Pfam; PF00037; Fer4; 1.
DR PRINTS; PR00353; 4FE4SFRDXIN.
DR TIGRFAMs; TIGR01945; rnfc; 1.
DR PROSITE; PS00198; 4FE4S_FEREDOXIN; 2.
KW 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding.
SQ SEQUENCE 449 AA; 49052 MW; 6F98DC0B98F52C78 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 25 HPPHG 29

RESULT 20
RHBA RHIME
ID RHBA RHIME STANDARD; PRT; 470 AA.
AC Q9Z3R2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Diaminobutyrate--2-oxoglutarate aminotransferase (EC 2.6.1.76) (L-
DE diaminobutyric acid transaminase) (Diaminobutyrate transaminase) (DABA
DE aminotransferase) (DABA-AT) (L-2,4-diaminobutyrate:2-ketoglutarate 4-
DE aminotransferase).
GN Name=rhba; Synonyms=rhsA; OrderedLocusNames=RA1258; ORFNames=SMA2400;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=21172875; PubMed=11274118;
RX DOI=10.1128/JB.183.8.2576-2585.2001;
RA Lynch D., O'Brien J., Welch T., Clarke P., Cuiv P.O., Crosa J.H.,
RA O'Connell M.;
RT "Genetic organization of the region encoding regulation, biosynthesis,
RT and transport of rhizobactin 1021, a siderophore produced by
RT Sinorhizobium meliloti.";
RL J. Bacteriol. 183:2576-2585(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432; DOI=10.1073/pnas.161294798;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hymen R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -|- CATALYTIC ACTIVITY: L-2,4-diaminobutyrate + 2-oxoglutarate = L-
CC glutamate + L-aspartic 4-semialdehyde.
CC -|- COFACTOR: Pyridoxal phosphate.
CC -|- PATHWAY: Siderophore biosynthesis; rhizobactin biosynthesis.
CC -|- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC aminotransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF110737; AAD09412.1; -.
DR EMBL; AE007311; AAK65916.1; -.
DR PIR; B95419; B95419.
DR PIR; T46814; T46814.
DR HSSP; P12995; 1QJ3.
DR InterPro; IPR005814; Aminotrans_3.
DR InterPro; IPR004637; Dat.
DR Pfam; PF00202; Aminotran_3; 1.
DR TIGRFAMs; TIGR00709; dat; 1.
DR TIGRFAMs; TIGR01885; Orn_aminotrans; 1.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
KW Aminotransferase; Complete proteome; Iron transport; Plasmid;
KW Pyridoxal phosphate; Transferase.

FT BINDING 304 304 Pyridoxal phosphate (Potential).
SQ SEQUENCE 470 AA; 50148 MW; 0377B3B8B9A09049 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 406 HPPHG 410

RESULT 21
Q73TJ7
ID Q73TJ7 PRELIMINARY; PRT; 491 AA.
AC Q73TJ7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3721;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017240; AAS06271.1; -.
KW Complete proteome.
SQ SEQUENCE 491 AA; 54266 MW; 442E24F9A831D8E0 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 11 HPPHG 15

RESULT 22
Q6P5Y4
ID Q6P5Y4 PRELIMINARY; PRT; 493 AA.
AC Q6P5Y4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC062581; AAH62581.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGc1; 2.
DR SMART; SM00408; IGc2; 4.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG_LIKE; 4.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 52761 MW; CD7C5591CC4822D2 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 477 HPPHG 481

RESULT 23
Q9W259
ID Q9W259 PRELIMINARY; PRT; 503 AA.
AC Q9W259;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG13506-PA.
GN ORFNames=CG13506;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosrin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA	Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,	
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA	Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,	
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	
RA	Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,	
RA	Williams S.M., Woodagat, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,	
RA	Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,	
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;	
RT	"The genome sequence of Drosophila melanogaster.";	
RL	Science 287:2185-2195(2000).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=22426065; PubMed=12537568;	
RA	Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,	
RA	Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,	
RA	George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,	
RA	Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,	
RA	Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,	
RA	Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;	
RT	"Finishing a whole-genome shotgun: Release 3 of the Drosophila	
RT	melanogaster euchromatic genome sequence.";	
RL	Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=22426070; PubMed=12537573;	
RA	Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,	
RA	Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,	
RA	Ashburner M., Celniker S.E.;	
RT	"The transposable elements of the Drosophila melanogaster euchromatin:	
RT	a genomics perspective.";	
RL	Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).	
RN	[4]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=22426069; PubMed=12537572;	
RA	Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,	
RA	Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,	
RA	Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,	
RA	Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,	
RA	Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,	
RA	Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,	
RA	Lewis S.E.;	
RT	"Annotation of the Drosophila melanogaster euchromatic genome: a	
RT	systematic review.";	
RL	Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).	
RN	[5]	
RP	SEQUENCE FROM N.A.	
RG	FlyBase;	
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.	
RN	[6]	
RP	SEQUENCE FROM N.A.	
RG	FlyBase;	
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AE003457; AAF46839.2; -.	
DR	FlyBase; FBgn0034723; CGI3506.	
DR	InterPro; IPR003599; Ig.	
DR	InterPro; IPR007110; Ig-like.	
DR	Pfam; PF00047; ig; 2.	
DR	SMART; SM00409; IG; 3.	
DR	PROSITE; PS50835; IG LIKE; 3.	
SQ	SEQUENCE 503 AA; 57072 MW; 8B37EE13309A511E CRC64;	
Query Match 100.0%; Score 36; DB 2; Length 503;		
Best Local Similarity 100.0%; Pred. No. 2.6e+02;		
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HPPHG 5	

Db	244 HPPHG 248	
RESULT 24		
HRG_RABIT	STANDARD; PRT; 526 AA.	
ID_Q28640;		
DT 01-NOV-1997 (Rel. 35, Created)		
DT 01-NOV-1997 (Rel. 35, Last sequence update)		
DT 05-JUL-2004 (Rel. 44, Last annotation update)		
DE Histidine-rich glycoprotein precursor (Histidine-proline rich		
DE glycoprotein) (HPRG) (Fragment).		
GN Name=HRG;		
OS Oryctolagus cuniculus (Rabbit).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX NCBI_TaxID=9986;		
RN [1]		
RP SEQUENCE FROM N.A., AND SEQUENCE OF 9-23; 301-313 AND 422-429.		
RC TISSUE=Serum;		
RX MEDLINE=96229917; PubMed=8639676; DOI=10.1021/bi952061t;		
RA Borza D.-B., Tatum F.M., Morgan W.T.;		
RT "Domain structure and conformation of histidine-proline-rich		
RT glycoprotein.";		
RL Biochemistry 35:1925-1934(1996).		
CC - - FUNCTION: The physiological function is not yet known. It binds		
CC heme, dyes and is known to interact with heparin, thrombospondin,		
CC and the lysine-binding site of plasminogen. On the basis of its		
CC homology with HMW kininogen, the His-rich region of this protein		
CC may mediate the contact activation phase of intrinsic blood		
CC coagulation cascade.		
CC - - SUBCELLULAR LOCATION: Secreted.		
CC - - TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.		
CC - - DOMAIN: In addition to having a high His and Pro content, this		
CC protein has many internal repeats. 15 tandem repetitions of a 5-		
CC residue sequence (G[H/P] [H/P]PH, consensus) form a His/Pro-rich		
CC region.		
CC - - SIMILARITY: Contains 2 cystatin-like domains.		
CC -----		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC the European Bioinformatics Institute. There are no restrictions on its		
CC use by non-profit institutions as long as its content is in no way		
CC modified and this statement is not removed. Usage by and for commercial		
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC or send an email to license@isb-sib.ch).		
CC -----		
CC EMBL; U32189; AAC48516.1; -.		
DR InterPro; IPR000010; Prot_inh_cystat.		
DR Pfam; PF00031; Cystatin; I.		
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat;		
KW Signal.		
FT NON TER 1 1		
FT SIGNAL <1 8		Potential.
FT CHAIN 9 526		Histidine-rich glycoprotein.
FT DOMAIN 127 243		Cystatin-like 1.
FT DOMAIN 251 296		Cystatin-like 2.
FT DOMAIN 329 498		Pro-rich.
FT DISULFID 14 505		His/Pro-rich.
FT DISULFID 68 79		By similarity.
FT DISULFID 95 116		By similarity.
FT DISULFID 193 415		By similarity.
FT DISULFID 207 230		By similarity.
FT DISULFID 272 302		Potential.
FT CARBOHYD 115 115		N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 192 192		N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 240 240		N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 310 310		N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 485 485		N-linked (GlcNAc. . .) (Potential).
FT SITE 303 304		Cleavage (by plasmin).

```
FT SITE 421 422 Cleavage (by plasmin).
SQ SEQUENCE 526 AA; 58877 MW; 810F23D367D93D42 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
   |||||
Db 353 HPPHG 357

RESULT 25
Q9YHE8
ID Q9YHE8 PRELIMINARY; PRT; 560 AA.
AC Q9YHE8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE HMG-box transcription factor (Transcription factor 7-like 1a) (T-cell
DE specific, HMG-box).
GN Name=tcf7l1a; Synonyms=TCF-3, tcf3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Pelegri F., Maischein H.-M.;
RT "Function of zebrafish beta-catenin and TCF-3 in dorsoventral
RT patterning.";
RL Mech. Dev. 77:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF097410; AAC64609.1; -.
DR EMBL; BC053135; AAHS3135.1; -.
DR HSSP; P27782; 2LEF.
DR ZFIN; ZDB-GENE-980605-30; tcf7l1a.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009071; HMG-box.
DR InterPro; IPR000910; HMG 12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS50118; HMG_BOX_2; 1.
```

```
SQ SEQUENCE 560 AA; 61459 MW; DC35C50EA4D61232 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
   |||||
Db 278 HPPHG 282

RESULT 26
Q87RZ6
ID Q87RZ6 PRELIMINARY; PRT; 581 AA.
AC Q87RZ6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VP0628.
GN OrderedLocusNames=VP0628;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005075; BAC58891.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR011550; Amidohydro_like.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_like; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 581 AA; 64511 MW; 292D0C34F2159F9E CRC64;

Query Match 100.0%; Score 36; DB 2; Length 581;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
   |||||
Db 348 HPPHG 352

RESULT 27
Q8D843
ID Q8D843 PRELIMINARY; PRT; 583 AA.
AC Q8D843;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted metal-dependent hydrolase with the TIM-barrel fold.
GN OrderedLocusNames=VV13145;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
```

DR	EMBL; AE016807; AA011465.1; --	OX	NCBI_TaxID=39947;
DR	GO; GO:0016787; F:hydrolase activity; IEA.	RN	[1]
DR	InterPro; IPR006680; Amidohydro 1.	RP	SEQUENCE FROM N.A.
DR	InterPro; IPR011550; Amidohydro_like.	RA	Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
DR	InterPro; IPR011059; Metalohydrolase.	RA	Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
DR	Pfam; PF01979; Amidohydro_1; 1.	RA	Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
DR	ProDom; PD001248; Amidohydro_like; 1.	RA	Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
KW	Complete proteome; Hydrolase.	RA	Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
SQ	SEQUENCE 583 AA; 64278 MW; 64AF798F8DCB69A6 CRC64;	RA	Wu H.-P., Shaw J.-F.;
Query Match	100.0%; Score 36; DB 2; Length 583;	Query Match	100.0%; Score 36; DB 2; Length 601;
Best Local Similarity	100.0%; Pred. No. 3e+02;	Best Local Similarity	100.0%; Pred. No. 3.1e+02;
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 HPPHG 5	Qy	1 HPPHG 5
Db	348 HPPHG 352	Db	5 HPPHG 9
RESULT 28		RESULT 30	
Q7MMD1		Q86BB8	
ID	Q7MMD1 PRELIMINARY; PRT; 587 AA.	ID	Q86BB8 PRELIMINARY; PRT; 723 AA.
AC	Q7MMD1;	AC	Q86BB8;
DT	01-MAR-2004 (TrEMBLrel. 26, Created)	DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)	DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Predicted metal-dependent hydrolase.	DE	CG30492-PD.
GN	OrderedLocusNames=VJ1141;	GN	ORFNames=CG30492;
OS	Vibrio vulnificus (strain VJ016).	OS	Drosophila melanogaster (Fruit fly).
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;	OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Vibrionaceae; Vibrio.	OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX	NCBI_TaxID=196600;	OC	Ephydroidea; Drosophilidae; Drosophila.
RN	[1]	OX	NCBI_TaxID=7227;
RP	SEQUENCE FROM N.A.	RN	[1]
RX	PubMed=14656965; DOI=10.1101/gr.1295503;	RP	SEQUENCE FROM N.A.
RA	Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,	RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA	Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,	RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;	RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RT	"Comparative genome analysis of Vibrio vulnificus, a marine	RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RT	pathogen.";	RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RL	Genome Res. 13:2577-2587(2003).	RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
DR	EMBL; AP005334; BAC93905.1; --	RA	Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
DR	GO; GO:0016787; F:hydrolase activity; IEA.	RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
DR	InterPro; IPR006680; Amidohydro 1.	RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
DR	InterPro; IPR011550; Amidohydro_like.	RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
DR	InterPro; IPR011059; Metalohydrolase.	RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
DR	Pfam; PF01979; Amidohydro_1; 1.	RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
DR	ProDom; PD001248; Amidohydro_like; 1.	RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
KW	Complete proteome; Hydrolase.	RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
SQ	SEQUENCE 587 AA; 64736 MW; B94C7FA5B3D15E7F CRC64;	RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Query Match	100.0%; Score 36; DB 2; Length 587;	Query Match	100.0%; Score 36; DB 2; Length 587;
Best Local Similarity	100.0%; Pred. No. 3e+02;	Best Local Similarity	100.0%; Pred. No. 3e+02;
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 HPPHG 5	Qy	1 HPPHG 5
Db	352 HPPHG 356	Db	352 HPPHG 356
RESULT 29		RESULT 29	
Q75HV8		Q75HV8	
ID	Q75HV8 PRELIMINARY; PRT; 601 AA.	ID	Q75HV8 PRELIMINARY; PRT; 601 AA.
AC	Q75HV8;	AC	Q75HV8;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Hypothetical protein OSJNBb0092G21.7.	DE	Hypothetical protein OSJNBb0092G21.7.
GN	Name=OSJNBb0092G21.7;	GN	Name=OSJNBb0092G21.7;
OS	Oryza sativa (japonica cultivar-group).	OS	Oryza sativa (japonica cultivar-group).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Ehrhartoideae; Oryzeae; Oryza.	OC	Ehrhartoideae; Oryzeae; Oryza.

RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AE003840; AA041497.1; -;
DR FlyBase; FBgn0050492; CG30492.
SQ SEQUENCE 723 AA; 78778 MW; C1D2AD1C04FF849C CRC64;

Query Match 100.0%; Score 36; DB 2; Length 723;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 514 HPPHG 518

RESULT 31
Q6C276 ID Q6C276 PRELIMINARY; PRT; 736 AA.
AC Q6C276;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to tr|Q873A0 Neurospora crassa B24N11.240.gene.
GN ORFNames=YALIOF10131g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dufon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekaiia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; CR382132; CAG78043.1; -;
SQ SEQUENCE 736 AA; 79795 MW; A63EAAE2BA4CE2A2 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 566 HPPHG 570

RESULT 32
Q95R83 ID Q95R83 PRELIMINARY; PRT; 773 AA.
AC Q95R83;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SD02660p.
GN Name=BcDNA:GH02712; ORFNames=CG30492;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY061574; AAL29122.1; -;
DR FlyBase; FBgn0050492; CG30492.
SQ SEQUENCE 773 AA; 84486 MW; 674466ECDC7B0FFF CRC64;

Query Match 100.0%; Score 36; DB 2; Length 773;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 564 HPPHG 568

RESULT 33

Q86BB9 PRELIMINARY; PRT; 938 AA.
AC Q86BB9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG30492-PE.
GN ORFNames=CG30492;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:

RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003840; AA041496.1; -.
DR FlyBase; FBgn0050492; CG30492.
DR InterPro; IPR011011; FYVE PHD ZnF.
DR InterPro; IPR011009; Kinase_like.
SQ SEQUENCE 938 AA; 102954 MW; 9984B6D579A1EA14 CRC64;
Query Match 100.0%; Score 36; DB 2; Length 938;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 729 HPPHG 733
RESULT 34
Q6XLI4
ID Q6XLI4 PRELIMINARY; PRT; 1075 AA.
AC Q6XLI4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Swan.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Y., Huang C.-H.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY226102; AAP48572.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; RRM_1; 3.
DR SMART; SM00360; RRM; 6.
DR PROSITE; PS50102; RRM; 3.
SQ SEQUENCE 1075 AA; 120486 MW; 24A993DB6F4BCBD8 CRC64;
Query Match 100.0%; Score 36; DB 2; Length 1075;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 880 HPPHG 884
RESULT 35
Q91562
ID Q91562 PRELIMINARY; PRT; 1427 AA.

AC Q91562;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tumor suppressor.
GN Name=XDCCa;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95113183; PubMed=7813784; DOI=10.1006/dbio.1994.1345;
RA Pierceall W.E., Reale M.A., Candia A.F., Wright C.V., Cho K.R.,
RA Fearon E.R.;
RT "Expression of a homologue of the deleted in colorectal cancer (DCC)
RT gene in the nervous system of developing Xenopus embryos.";
RL Dev. Biol. 166:654-665(1994).
DR EMBL; U10986; AAA70168.1; -.
DR PIR; I51669; I51669.
DR HSSP; Q9UQH9; 1EV2.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF06583; Neogenin_C; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50853; FN3; 6.
DR PROSITE; PS50835; IG_LIKE; 4.
SQ SEQUENCE 1427 AA; 156533 MW; 61FEA12C8A674972 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 1427;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 1066 HPPHG 1070

RESULT 36
Q63155 PRELIMINARY; PRT; 1445 AA.
AC Q63155;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Deleted in Colorectal Cancer.
GN Name=DCC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97015074; PubMed=8861902; DOI=10.1016/S0092-8674(00)81336-7;
RA Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.-Y.,
RA Culotti J.G., Tessier-Lavigne M.;
RT "Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RL Cell 87:175-185(1996).
DR EMBL; U68725; AAB41099.1; -.
DR HSSP; P13596; 1QZ1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig c2.
DR InterPro; IPR010560; Neogenin_C.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF06583; Neogenin_C; 1.
DR PRINTS; PR00014; FNTYPEIII.

DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50853; FN3; 6.
DR PROSITE; PS50835; IG_LIKE; 4.
SQ SEQUENCE 1445 AA; 157940 MW; 084F625954481988 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 1445;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 37
DCC_HUMAN STANDARD; PRT; 1447 AA.
ID _DCC_HUMAN
AC P43146;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Netrin receptor DCC precursor (tumor suppressor protein DCC)
DE (Colorectal cancer suppressor).
GN Name=DCC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011532; PubMed=7926722;
RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,
RA Vogelstein B.;
RT "The DCC gene product in cellular differentiation and colorectal
RT tumorigenesis.";
RL Genes Dev. 8:1174-1183(1994).
RN [2]
RP SEQUENCE OF 1-750 FROM N.A.
RX MEDLINE=90100559; PubMed=2294591;
RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,
RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
RA Vogelstein B.;
RT "Identification of a chromosome 18q gene that is altered in colorectal
RT cancers.";
RL Science 247:49-56(1990).
RN [3]
RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBELD EXONS).
RX MEDLINE=91121517; PubMed=1991322; DOI=10.1016/0092-8674(91)90244-S;
RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,
RA Oliner J.D., Kinzler K.W., Vogelstein B.;
RT "Scrambled exons.";
RL Cell 64:607-613(1991).
RN [4]
RP GENE STRUCTURE, AND VARIANT CARCINOMA HIS-1375.
RX MEDLINE=94245241; PubMed=8188295;
RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
RT "The DCC gene: structural analysis and mutations in colorectal
RT carcinomas.";
RL Genomics 19:525-531(1994).
RN [5]
RP FUNCTION.
RX MEDLINE=97015074; PubMed=8861902; DOI=10.1016/S0092-8674(00)81336-7;
RA Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.,
RA Culotti J.G., Tessier-Lavigne M.;
RT "Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RL Cell 87:175-185(1996).
RN [6]
RP INTERACTIONS WITH SIAH1 AND SIAH2, AND DEGRADATION.
RX PubMed=9334332;
RA Hu G., Zhang S., Vidal M., Baer J.L., Xu T., Fearon E.R.;
RT "Mammalian homologs of seven in absentia regulate DCC via the

RT ubiquitin-proteasome pathway.";

RL Genes Dev. 11:2701-2714(1997).

RN [7]

RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.

RX MEDLINE=94243823; PubMed=8187090;

RA Miyake S., Nagai K., Yoshino K., Oto M., Endo M., Yuasa Y.;

RT "Point mutations and allelic deletion of tumor suppressor gene DCC in

RT human esophageal squamous cell carcinomas and their relation to

RT metastasis.";

RL Cancer Res. 54:3007-3010(1994).

CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates

CC axon attraction of neuronal growth cones in the developing nervous

CC system upon ligand binding. Its association with UNC5 proteins may

CC trigger signaling for axon repulsion. It also acts as a dependence

CC receptor required for apoptosis induction when not associated with

CC netrin ligand. Implicated as a tumor suppressor gene.

CC -!- SUBUNIT: Interacts with the cytoplasmic part of UNC5A, UNC5B,

CC UNC5C and probably UNC5D (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Found in axons of the central and peripheral

CC nervous system and in differentiated cell types of the intestine.

CC -!- PTM: Ubiquitinated; mediated by SIAH1 or SIAH2 and leading to its

CC subsequent proteasomal degradation (Probable).

CC -!- DISEASE: Colorectal tumors that lost their capacity to

CC differentiate into mucus producing cells uniformly lack DCC

CC expression. Inactivation of DCC due to allelic deletion and/or

CC point mutations may cause both lymphatic and hematogenous

CC metastasis of esophageal squamous cell carcinomas.

CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. DCC family.

CC -!- SIMILARITY: Contains 6 fibronectin type III domains.

CC -!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; X76132; CAA53735.1; -.

DR EMBL; M32292; AAA35751.1; -.

DR EMBL; M32286; AAA52174.1; -.

DR EMBL; M32288; AAA52175.1; ALT_SEQ.

DR EMBL; M32290; AAA52176.1; -.

DR EMBL; M63696; AAA52177.1; -.

DR EMBL; M63700; AAA52178.1; -.

DR EMBL; M63702; AAA52179.1; -.

DR EMBL; M63718; AAA52180.1; -.

DR EMBL; M63698; AAA52181.1; -.

DR PIR; A54100; A54100.

DR HSSP; P13590; 1IE5.

DR Genew; HGNC:2701; DCC.

DR MIM; 120470; -.

DR GO; GO:0004888; F:transmembrane receptor activity; TAS.

DR GO; GO:0007409; P:axonogenesis; TAS.

DR GO; GO:0006917; P:induction of apoptosis; TAS.

DR InterPro; IPR003961; FN_III.

DR InterPro; IPR008957; FN_III-like.

DR InterPro; IPR003962; FN_III subd.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR010560; Neogenin_C.

DR Pfam; PF00041; fn3; 6.

DR Pfam; PF00047; ig; 4.

DR Pfam; PF06583; Neogenin_C; 1.

DR PRINTS; PR00014; FNTYPEIII.

DR SMART; SM00060; FN3; 6.

DR SMART; SM00408; IGC2; 3.

DR PROSITE; PS50853; FN3; 6.

DR PROSITE; PS50835; IG_LIKE; 4.

KW Anti-oncogene; Apoptosis; Developmental protein; Disease mutation;

KW Glycoprotein; Immunoglobulin domain; Polymorphism; Receptor; Repeat;

KW Signal; Transmembrane; Ubl conjugation.

FT SIGNAL 1 25 Potential.

FT CHAIN 26 1447 Netrin receptor DCC.

FT DOMAIN 26 1097 Extracellular (Potential).

FT TRANSMEM 1098 1122 Potential.

FT DOMAIN 1123 1447 Cytoplasmic (Potential).

FT DOMAIN 26 135 Ig-like C2-type 1.

FT DOMAIN 139 229 Ig-like C2-type 2.

FT DOMAIN 234 326 Ig-like C2-type 3.

FT DOMAIN 331 416 Ig-like C2-type 4.

FT DOMAIN 429 520 Fibronectin type-III 1.

FT DOMAIN 528 616 Fibronectin type-III 2.

FT DOMAIN 622 714 Fibronectin type-III 3.

FT DOMAIN 726 814 Fibronectin type-III 4.

FT DOMAIN 843 939 Fibronectin type-III 5.

FT DOMAIN 944 1041 Fibronectin type-III 6.

FT DISULFID 61 117 By similarity.

FT DISULFID 161 212 By similarity.

FT DISULFID 261 310 By similarity.

FT DISULFID 352 400 By similarity.

FT CARBOHYD 94 94 N-linked (GlcNAc . .) (Potential).

FT CARBOHYD 299 299 N-linked (GlcNAc . .) (Potential).

FT CARBOHYD 318 318 N-linked (GlcNAc . .) (Potential).

FT CARBOHYD 478 478 N-linked (GlcNAc . .) (Potential).

FT CARBOHYD 628 628 N-linked (GlcNAc . .) (Potential).

FT CARBOHYD 702 702 N-linked (GlcNAc . .) (Potential).

FT VARIANT 168 168 M -> T (in a esophageal carcinoma).

FT VARIANT 201 201 /FTID=VAR 003909.

FT VARIANT 1375 1375 R -> G (in dbSNP:2229080).

FT VARIANT 1375 1375 P -> H (in a colorectal carcinoma).

FT CONFLICT 138 138 /FTID=VAR 003911.

FT CONFLICT 233 329 Missing (in Ref. 3).

FT CONFLICT 421 421 Missing (in Ref. 3).

SQ SEQUENCE 1447 AA; 158456 MW; 4A8612766ED0471F CRC64;

Query Match 100.0%; Score 36; DB 1; Length 1447;

Best Local Similarity 100.0%; Pred. No. 7.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5

Db 1084 HPPHG 1088

RESULT 38

DCC_MOUSE

ID DCC_MOUSE STANDARD; PRT; 1447 AA.

AC P70211;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Netrin receptor DCC precursor (Tumor suppressor protein DCC).

GN Name=Dcc;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).

RC STRAIN=BALB/c; TISSUE=Brain;

RX MEDLINE=96112625; PubMed=8570174;

RA Cooper H.M., Armes P., Britto J., Gad J., Wilks A.F.;

RT "Cloning of the mouse homologue of the deleted in colorectal cancer

RL gene (mDCC) and its expression in the developing mouse embryo.";

RN Oncogene 11:2243-2254(1995).

RN [2]

RP REVISIONS.

RC STRAIN=BALB/c; TISSUE=Brain;

RA Cooper H.M.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates

axon attraction of neuronal growth cones in the developing nervous system upon ligand binding. Its association with UNC5 proteins may trigger signaling for axon repulsion. It also acts as a dependence receptor required for apoptosis induction when not associated with netrin ligand. Implicated as a tumor suppressor gene.

-!- SUBUNIT: Interacts with the cytoplasmic part of UNC5A, UNC5B, UNC5C and probably UNC5D (By similarity).

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=2;

Name=A;

Isold=P70211-1; Sequence=Displayed;

Note=Isoform B is produced by alternative initiation at Met-85 of isoform A;

Name=C;

Isold=P70211-2; Sequence=VSP_002501;

Event=Alternative initiation;

Comment=2 isoforms, A (shown here) and B, are produced by alternative initiation at Met-1 and Met-85;

-!- TISSUE SPECIFICITY: In the embryo, expressed at high levels in the developing brain and neural tube. In adult, highly expressed in brain with very low levels found in testis, heart and thymus.

-!- Isoform C is expressed only in the embryo.

-!- DEVELOPMENTAL STAGE: Low levels in early gestation. Highest levels expressed during mid gestation. Levels decrease in late gestation and remain at this level in the adult.

-!- PTM: Ubiquitinated; mediated by SIAH1 or SIAH2 and leading to its subsequent proteasomal degradation (By similarity).

-!- SIMILARITY: Belongs to the immunoglobulin superfamily. DCC family.

-!- SIMILARITY: Contains 6 fibronectin type III domains.

-!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X85788; CAA59786.1; -.
HSSP; P13590; 1IE5.
MGD; MGI:94869; Dcc.
InterPro; IPR003961; FN_III.
InterPro; IPR008957; FN_III-like.
InterPro; IPR003962; FN_III subd.
InterPro; IPR007110; Ig-like.
InterPro; IPR003598; Ig_C2.
InterPro; IPR010560; Neogenin_C.
Pfam; PF0041; fn3; 6.
Pfam; PF00047; ig; 4.
Pfam; PF06583; Neogenin_C; 1.
PRINTS; PR00014; FNTYPEII.
SMART; SM00060; FN3; 6.
SMART; SM00408; IGC2; 3.
PROSITE; PS50853; FN3; 6.
PROSITE; PS50835; IG_LIKE; 4.
Alternative initiation; Alternative splicing; Anti-oncogene;
Apoptosis; Developmental protein; Glycoprotein; Immunoglobulin domain;
Receptor; Repeat; Signal; Transmembrane; Ubl conjugation.

SIGNAL 1 25 Potential.
CHAIN 26 1447 Netrin receptor DCC, isoform A.
CHAIN 85 1447 Netrin receptor DCC, isoform B.
INIT_MET 85 85 For isoform B.
DOMAIN 26 1097 Extracellular (Potential).
TRANSMEM 1098 1122 Potential.
DOMAIN 1123 1447 Cytoplasmic (Potential).
DOMAIN 36 135 Ig-like C2-type 1.
DOMAIN 139 229 Ig-like C2-type 2.
DOMAIN 234 326 Ig-like C2-type 3.
DOMAIN 331 416 Ig-like C2-type 4.
DOMAIN 429 520 Fibronectin type-III 1.
DOMAIN 528 616 Fibronectin type-III 2.

FT	DOMAIN	622	714	Fibronectin type-III 3.
FT	DOMAIN	726	814	Fibronectin type-III 4.
FT	DOMAIN	843	939	Fibronectin type-III 5.
FT	DOMAIN	944	1041	Fibronectin type-III 6.
FT	DISULFID	61	117	By similarity.
FT	DISULFID	161	212	By similarity.
FT	DISULFID	261	310	By similarity.
FT	DISULFID	352	400	By similarity.
FT	CARBOHYD	60	60	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	94	94	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	299	299	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	318	318	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	478	478	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	628	628	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	702	702	N-linked (GlcNAc. . .) (Potential).
FT	VARSPPLIC	819	838	Missing (in isoform C).
FT				/FTid=VSP_002501.
SQ	SEQUENCE	1447 AA;	158298 MW;	OD1F1097C22D5B9F CRC64;

Query Match 100.0%; Score 36; DB 1; Length 1447;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 HPPHG 5
Db	1084 HPPHG 1088

RESULT 39

Q6SLE5

ID Q6SLE5 PRELIMINARY; PRT; 2013 AA.

AC Q6SLE5;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Putative histidine kinase HHK2p.

GN Name=HHK2;

OS Cochliobolus heterostrophus (Drechslera maydis).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;

OC Pleosporales; Pleosporaceae; Cochliobolus.

OX NCBI_TaxID=5016;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C4;

RX PubMed=14665450; DOI=10.1128/EC.2.6.1151-1161.2003;

RA Catlett N.L., Yoder O.C., Turgeon B.G.;

RT "Whole-genome analysis of two-component signal transduction genes in fungal pathogens.";

RT Eukaryotic Cell 2:1151-1161(2003).

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

DR EMBL; AY456005; AAR29881.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR011006; CheY_like.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000700; PAS-assoc_C.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HisKA_1.
DR Pfam; PF00785; PAC; 2.
DR Pfam; PF00989; PAS; 2.
DR Pfam; PF00072; Response_reg; 1.

DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00086; PAC; 2.
DR SMART; SM00091; PAS; 2.
DR SMART; SM00448; REC; 1.
DR TIGRFAMS; TIGR00229; sensory_box; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50113; PAC; 2.
DR PROSITE; PS50112; PAS; 2.
DR PROSITE; PS50110; RESPONSE REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction.
SQ SEQUENCE 2013 AA; 219997 MW; E4B99CFCC9D57065 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 2013;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 94 HPPHG 98

RESULT 40
Q9S8A0
ID Q9S8A0 PRELIMINARY; PRT; 16 AA.
AC Q9S8A0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Ribulose BIPHOSPHATE carboxylase small subunit (Fragment).
OS Pinus monticola (Western white pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3345;
RN [1]
RP SEQUENCE.
RX MEDLINE=96213005; PubMed=8665095;
RA Ekramoddoullah A.K., Taylor D.W.;
RT "Seasonal variation of western white pine (Pinus monticola D. Don)
RT foliage proteins.";
RL Plant Cell Physiol. 37:189-199(1996).
SQ SEQUENCE 16 AA; 1803 MW; 27D4934C881717DE CRC64;

Query Match 83.3%; Score 30; DB 2; Length 16;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
Db 4 HPPYG 8

Search completed: June 15, 2005, 14:21:27
Job time : 86.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-10
Perfect score: 35
Sequence: 1 PPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	5	5	ABB79811 Histidine
2	35	100.0	5	8	ADH10416 Rabbit HP
3	35	100.0	17	3	AAB39278 Human sec
4	35	100.0	53	4	ABB70178 Drosophil
5	35	100.0	55	4	AAM15229 Peptide #
6	35	100.0	55	4	ABB34219 Peptide #
7	35	100.0	55	4	AAM27690 Peptide #
8	35	100.0	55	4	ABB29058 Peptide #
9	35	100.0	55	4	ABB19655 Protein #
10	35	100.0	55	4	AAM67398 Human bon
11	35	100.0	55	4	AAM55013 Human bra
12	35	100.0	55	4	ABG49057 Human liv
13	35	100.0	55	4	AAM02972 Peptide #
14	35	100.0	55	5	ABG37025 Human pep
15	35	100.0	57	4	AAU54664 Propionib
16	35	100.0	57	6	ABM51183 Propionib
17	35	100.0	60	4	AAU43160 Propionib
18	35	100.0	60	6	ABM39679 Propionib
19	35	100.0	61	4	AAM37888 Peptide #
20	35	100.0	61	4	AAM77682 Human bon
21	35	100.0	61	4	AAM64953 Human bra
22	35	100.0	61	4	ABG59329 Human liv
23	35	100.0	75	4	AAU32380 Novel hum
24	35	100.0	77	4	AAU41116 Propionib
25	35	100.0	77	4	AAU43934 Propionib

26	35	100.0	77	6	ABM40453	Abm40453 Propionib
27	35	100.0	77	6	ABM37635	Abm37635 Propionib
28	35	100.0	79	2	AAy12827	Aay12827 Human 5'
29	35	100.0	80	4	ABB61183	Abb61183 Drosophil
30	35	100.0	89	5	ABP11024	Abp11024 Human ORF
31	35	100.0	89	6	ABU34317	Abu34317 Protein e
32	35	100.0	97	4	ABG24494	Abg24494 Novel hum
33	35	100.0	98	7	ABO63515	Abo63515 Klebsiell
34	35	100.0	101	8	ADH10412	Adh10412 Rabbit HP
35	35	100.0	103	7	ABO66179	Abo66179 Klebsiell
36	35	100.0	111	4	AAG67890	Aag67890 Human gly
37	35	100.0	111	4	AAG64349	Aag64349 Glycosyl
38	35	100.0	115	3	AAG16968	Aag16968 Arabidops
39	35	100.0	116	3	AAG16967	Aag16967 Arabidops
40	35	100.0	118	4	AAO01973	Aao01973 Human pol
41	35	100.0	124	4	AAO04253	Aao04253 Human pol
42	35	100.0	124	7	ADC32845	Adc32845 Human nov
43	35	100.0	124	7	ADF60519	Adf60519 Human con
44	35	100.0	129	4	AAU44880	Aau44880 Propionib
45	35	100.0	129	6	ABM41399	Abm41399 Propionib

ALIGNMENTS

RESULT 1
ABB79811
ID ABB79811 standard; peptide; 5 AA.

XX AC ABB79811;
XX XX
DT 25-NOV-2002 (first entry)
XX XX
DE Histidine proline rich glycoprotein pentapeptide.
XX XX
KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW nootropic; neuroprotective; antiparkinsonian.
XX XX
OS Synthetic.
XX XX
PN WO200264621-A2.
XX XX
PD 22-AUG-2002.
XX XX
PF 14-FEB-2002; 2002WO-US004336.
XX XX
PR 14-FEB-2001; 2001US-0268370P.
XX XX
PA (ATTE-) ATTENUON LLC.
XX XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX XX
DR WPI; 2002-666989/71.
XX XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX XX
PS Claim 2; Page 67; 82pp; English.

The present sequence is a specific example of claimed anti-angiogenic pentapeptides of the invention. Claimed anti-angiogenic polypeptides or peptides comprise: the histidine-proline-rich (H/P) domain of human histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain of rabbit HPRG (see ABB79807); a variant of these that is capable of inhibiting angiogenesis, endothelial cell proliferation or endothelial tube formation in vitro or in vivo; or a pentapeptide having the generic sequence given in ABB79808, such as the present peptide, or its variant having an additional 1 to 4 amino acids comprising His, Pro or Gly at its N- or C-terminus. Also claimed are: chemically synthesised or recombinantly produced peptide multimers; a diagnostically or

CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays

XX Sequence 5 AA;

Query Match 100.0%; Score 35; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 1 PPPHG 5

RESULT 2
ADH10416
ID ADH10416 standard; peptide; 5 AA.
XX
AC ADH10416;

XX 11-MAR-2004 (first entry)

DE Rabbit HPRG protein H/P rich domain repeat fragment.

XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.

OS Oryctolagus cuniculus.

XX WO2003077872-A2.

PN 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

PF 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

XX

PI Mcrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Example 5; SEQ ID NO 30; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

Query Match 100.0%; Score 35; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 1 PPPHG 5

RESULT 3
AAB39278
ID AAB39278 standard; protein; 17 AA.
XX
AC AAB39278;

XX 02-FEB-2001 (first entry)

DE Human secreted protein sequence encoded by gene 31 SEQ ID NO:158.

XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide; neoplasm;
KW ophthalmological; autoimmune disease; rheumatoid arthritis; angiogenesis;
KW hyperproliferative disorder; cardiovascular disorder; infection;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; chemotaxis.

OS Homo sapiens.

XX WO200056754-A1.

XX 28-SEP-2000.

XX

PF 16-MAR-2000; 2000WO-US006792.
XX
PR 19-MAR-1999; 99US-0125362P.
PR 10-DEC-1999; 99US-0169980P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen GA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-579483/54.
DR N-PSDB; AAC74253.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Disclosure; Page 52; 434pp; English.
XX
CC The polynucleotide sequences given in AAC74223-C74279 encode the human
CC secreted proteins represented in AAB39179-B39226. Sequences AAB39227-
CC B39308 are alternative proteins encoded by the genes, and also protein
CC sequences with which they share homology. The proteins have activities
CC based on the tissues and cells in which they are expressed. Examples of
CC activities include: immunosuppressive; antiarthritic; antirheumatic;
CC antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;
CC neurotropic; neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The human secreted proteins, polynucleotides,
CC antagonists and agonists of the invention may be useful in the treatment,
CC prevention, and/or diagnosis of various disease, disorders and conditions
CC such as autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC e.g. corneal infection. The polypeptides can also be used to aid wound
CC healing and epithelial cell proliferation, to regenerate tissues,
CC maintain organs before transplantation, in chemotaxis and as a food
CC additive or preservative e.g. to increase storage capabilities. Sequences
CC AAC74214-C74222 and AAB39178 are used during the isolation and
CC characterisation of the genes of the invention
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 35; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db |||||
2 PPPHG 6

RESULT 4
AAB70178
ID ABB70178 standard; protein; 53 AA.
XX
AC ABB70178;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 37326.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL14281.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 37326; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 35; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db |||||
2 PPPHG 6

RESULT 5
AAM15229
ID AAM15229 standard; protein; 55 AA.
XX
AC AAM15229;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #1663 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human cervical epithelial cells.
XX
PS Claim 27; SEQ ID NO 20055; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
1 PPPHG 5

RESULT 6
ABB34219
ID ABB34219 standard; peptide; 55 AA.
XX
AC ABB34219;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #1725 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX
PS Claim 27; SEQ ID NO 26854; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
1 PPPHG 5

RESULT 7
AAM27690
ID AAM27690 standard; protein; 55 AA.
XX
AC AAM27690;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #1727 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 27959; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENPs:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
1 PPPHG 5

RESULT 8
ABB29058
ID ABB29058 standard; peptide; 55 AA.

XX ABB29058;
AC
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #1709 encoded by breast cell single exon nucleic acid probe.
XX
DE Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
KW
XX Homo sapiens.
OS
XX WO200157271-A2.
PN
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US0000662.
PF
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
DR
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 12026; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
1 PPPHG 5

RESULT 9
ABB19655
ID ABB19655 standard; protein; 55 AA.
XX
AC ABB19655;

XX 23-JAN-2002 (first entry)
DT
XX
DE Protein #1654 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
XX Homo sapiens.
OS
XX WO200157274-A2.
PN
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US0000666.
PF
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
PT
XX
PS Claim 15; SEQ ID NO 21425; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
1 PPPHG 5

RESULT 10
AAM67398
ID AAM67398 standard; protein; 55 AA.
XX
AC AAM67398;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27704.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX

OS Homo sapiens.
XX WO200157276-A2.
PN
XX
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow.
PT
PT
XX
XX Example 4; SEQ ID NO 27704; 658pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is a protein encoded by one of the probes of the invention
CC
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHHG 5
Db |||||
1 PPHHG 5

RESULT 11
AAM55013
ID AAM55013 standard; protein; 55 AA.
XX
XX AAM55013;
AC
XX 05-NOV-2001 (first entry)
DT
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 27118.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray; Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
KW
KW Homo sapiens.
XX
OS
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000667.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-483446/52.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human brains.
PT
XX Example 4; SEQ ID NO 27118; 650pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention
CC
XX Sequence 55 AA;
SQ

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHHG 5
Db |||||
1 PPHHG 5

RESULT 12
ABG49057
ID ABG49057 standard; peptide; 55 AA.
XX
XX ABG49057;
AC
XX 25-FEB-2003 (first entry)
DT
XX Human liver peptide, SEQ ID NO 27705.
DE
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease.
KW
KW Homo sapiens.
XX
OS
XX WO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000664.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488898/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human adult liver.
PT
XX Claim 27; SEQ ID NO 27705; 658pp; English.
PS
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 1 PPPHG 5

RESULT 13
AAM02972
ID AAM02972 standard; protein; 55 AA.
XX
AC AAM02972;

DT 09-OCT-2001 (first entry)

XX Peptide #1654 encoded by probe for measuring breast gene expression.
DE Probe; human; breast disease; breast cancer; development disorder;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
KW
KW
XX

OS Homo sapiens.

XX WO200157270-A2.

PN
XX
PD 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US000661.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 27; SEQ ID NO 11712; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes
CC (see AA100010-AA110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 1 PPPHG 5

RESULT 14
ABG37025
ID ABG37025 standard; peptide; 55 AA.
XX
AC ABG37025;

DT 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 26690.
DE
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 26690; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 5; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 1 PPPHG 5

RESULT 15
AAU54664
ID AAU54664 standard; protein; 57 AA.
XX
AC AAU54664;
XX

DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #15560.
XX

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS595566.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 15859; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 57 AA;

Query Match 100.0%; Score 35; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 47 PPPHG 51

RESULT 16
ABM51183
ID ABM51183 standard; protein; 57 AA.
XX
AC ABM51183;
XX

DT 20-OCT-2003 (first entry)
XX

DE Propionibacterium acnes predicted ORF-encoded polypeptide #15859.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO2003033515-A1.

PD 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

PF 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglas J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64495.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 15859; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 57 AA;

Query Match 100.0%; Score 35; DB 6; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 47 PPPHG 51

RESULT 17
AAU43160
ID AAU43160 standard; protein; 60 AA.
XX
AC AAU43160;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #4056.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-020841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59519.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 4355; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 AA;

Query Match 100.0%; Score 35; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 46 PPPHG 50

RESULT 18
ABM39679
ID ABM39679 standard; protein; 60 AA.
XX
AC ABM39679;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #4355.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglas J;
XX

DR WPI; 2003-381789/36.
DR N-PSDB; ACF64448.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 4355; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
CC via this method; a vaccine composition (comprising P. acnes polypeptides, or
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 AA;

Query Match 100.0%; Score 35; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db |||||
46 PPHG 50

RESULT 19
AAM37888
ID AAM37888 standard; protein; 61 AA.
XX
AC AAM37888;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #11925 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 38157; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 61 AA;

Query Match 100.0%; Score 35; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db |||||
18 PPHG 22

RESULT 20
AAM77682
ID AAM77682 standard; protein; 61 AA.
XX
AC AAM77682;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 37988.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 37988; 658pp + Sequence Listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 61 AA;

Query Match 100.0%; Score 35; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 18 PPPHG 22

RESULT 21
AAM64953
ID AAM64953 standard; protein; 61 AA.

XX
AC AAM64953;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 37058.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX

OS Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX Example 4; SEQ ID NO 37058; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention

XX Sequence 61 AA;

Query Match 100.0%; Score 35; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 18 PPPHG 22

RESULT 22
ABG59329

ID ABG59329 standard; peptide; 61 AA.

XX
AC ABG59329;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 37977.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX

OS Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000664.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX Claim 27; SEQ ID NO 37977; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG59930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 61 AA;

Query Match 100.0%; Score 35; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 18 PPPHG 22

RESULT 23
AAU32380

ID AAU32380 standard; protein; 75 AA.

XX AAU32380;
AC
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #2871.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 597-598; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 75 AA;

Query Match 100.0%; Score 35; DB 4; Length 75;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 3 PPPHG 7

RESULT 24
AAU41116
ID AAU41116 standard; protein; 77 AA.
XX
AC AAU41116;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #2012.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-020841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59514.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 2311; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 77 AA;

Query Match 100.0%; Score 35; DB 4; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 27 PPPHG 31

RESULT 25
AAU43934
ID AAU43934 standard; protein; 77 AA.
XX
AC AAU43934;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #4830.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
OS Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX (CORI-) CORIXA CORP.
PA
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59521.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 5129; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 77 AA;
Query Match 100.0%; Score 35; DB 4; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db 13 PPPHG 17
RESULT 26
ABM40453
ID ABM40453 standard; protein; 77 AA.
XX
AC ABM40453;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #5129.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX 24-APR-2003.
XX 11-OCT-2002; 2002WO-US032727.
PF
XX 15-OCT-2001; 2001US-00978825.
PR (CORI-) CORIXA CORP.
XX
PA Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglas J;
PI
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64450.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 5129; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC polynucleotide of the invention comprising a polypeptide of the invention; a
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 77 AA;
Query Match 100.0%; Score 35; DB 6; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db 13 PPPHG 17
RESULT 27
ABM37635
ID ABM37635 standard; protein; 77 AA.
XX
AC ABM37635;
XX
DT 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #2311.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64443.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 2311; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 77 AA;

Query Match 100.0%; Score 35; DB 6; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 27 PPPHG 31

RESULT 28
AAY12827
ID AAY12827 standard; protein; 79 AA.

XX AAY12827;
AC
XX
DT 21-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO:417.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9906549-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001231.
XX
PR 01-AUG-1997; 97US-00905279.
XX (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153779/13.
DR N-PSDB; AAX51605.
XX
PT New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries derived from testis, ovary, uterus and spleen tissue.
XX
PS Claim 34; Page 469; 522pp; English.
XX
CC AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12681 to
CC AAY12913, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 79 AA;

Query Match 100.0%; Score 35; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 59 PPPHG 63

RESULT 29
ABB61183
ID ABB61183 standard; protein; 80 AA.
XX
AC ABB61183;
XX
DT 26-MAR-2002 (first entry)
XX

PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA38187.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 62241; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 89 AA;

Query Match 100.0%; Score 35; DB 6; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|
Db 82 PPPHG 86

RESULT 32
ABG24494
ID ABG24494 standard; protein; 97 AA.
XX
AC ABG24494;
XX

DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24485.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88681.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 54853; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 97 AA;

Query Match 100.0%; Score 35; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|
Db 51 PPPHG 55

RESULT 33
ABO63515
ID ABO63515 standard; protein; 98 AA.
XX
AC ABO63515;
XX
DT 29-JUL-2004 (first entry)
XX

DE Klebsiella pneumoniae polypeptide seqid 10032.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
PF
XX 29-JAN-1999; 99US-0117747P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Breton GL, Osborne M;
PI
XX WPI; 2003-895346/82.
DR
DR N-PSDB; ACH97066.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
PT
XX Disclosure; SEQ ID NO 10032; 932pp; English.
PS
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
SQ Sequence 98 AA;

Query Match 100.0%; Score 35; DB 7; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 28 PPPHG 32

RESULT 34
ADH10412
ID ADH10412 standard; protein; 101 AA.
XX
AC ADH10412;
XX
DT 11-MAR-2004 (first entry)
XX
DE Rabbit HPRG protein H/P rich domain.
XX
KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiac; vasotropic; vulneryary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Oryctolagus cuniculus.
XX
PN WO2003077872-A2.
XX
PD 25-SEP-2003.
XX
XX 17-MAR-2003; 2003WO-US008060.
PF
XX 15-MAR-2002; 2002US-0364047P.
PR
XX (ATTE-) ATTENUON LLC.
PA

XX McCrae K, Donate F, Juarez J, Mazar AP;
PI
XX WPI; 2004-090604/09.
DR
XX
PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
PS Claim 6; SEQ ID NO 24; 117pp; English.
XX
CC The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 35; DB 8; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 8 PPPHG 12

RESULT 35
ABO66179
ID ABO66179 standard; protein; 103 AA.
XX
AC ABO66179;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 12696.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
PF
XX 29-JAN-1999; 99US-0117747P.
PR
XX (ATTE-) ATTENUON LLC.
PA

XX AC AAG16968;
XX XX
DT 17-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17811.
XX XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX XX
OS Arabidopsis thaliana.
XX XX
PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-00301439.
XX XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 35; DB 3; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 79 PPPHG 83

RESULT 39
AAG16967
ID AAG16967 standard; protein; 116 AA.

XX AC AAG16967;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 17810.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 35; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 80 PPPHG 84

RESULT 40
AAO01973
ID AAO01973 standard; protein; 118 AA.

XX AAO01973;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 15865.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX

PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI81904.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 15865; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 118 AA;

Query Match 100.0%; Score 35; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
18 PPPHG 22

Search completed: June 15, 2005, 14:15:52
Job time : 92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-10
Perfect score: 35
Sequence: 1 PPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTus_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	35	100.0	52	1 US-07-745-206A-28	Sequence 28, Appl
2	35	100.0	52	2 US-08-311-363-28	Sequence 28, Appl
3	35	100.0	98	4 US-09-489-039A-10032	Sequence 10032, A
4	35	100.0	103	4 US-09-489-039A-12696	Sequence 12696, A
5	35	100.0	135	4 US-09-270-767-32232	Sequence 32232, A
6	35	100.0	135	4 US-09-270-767-47449	Sequence 47449, A
7	35	100.0	174	2 US-08-683-262B-47	Sequence 47, Appl
8	35	100.0	174	3 US-09-361-707-47	Sequence 47, Appl
9	35	100.0	202	4 US-09-640-211A-1165	Sequence 1165, Ap
10	35	100.0	240	5 PCT-US93-00227-4	Sequence 4, Appli
11	35	100.0	267	4 US-09-270-767-43608	Sequence 43608, A
12	35	100.0	321	4 US-09-252-991A-32545	Sequence 32545, A
13	35	100.0	332	4 US-09-270-767-45004	Sequence 45004, A
14	35	100.0	335	4 US-08-894-139-4	Sequence 4, Appli
15	35	100.0	354	4 US-09-949-016-11268	Sequence 11268, A
16	35	100.0	434	4 US-09-949-016-6701	Sequence 6701, Ap
17	35	100.0	439	4 US-09-248-796A-18484	Sequence 18484, A
18	35	100.0	479	2 US-08-149-097D-38	Sequence 38, Appl
19	35	100.0	491	4 US-10-029-180-106	Sequence 106, App
20	35	100.0	523	1 US-08-455-543A-42	Sequence 42, Appl
21	35	100.0	523	2 US-08-223-305C-42	Sequence 42, Appl
22	35	100.0	524	1 US-08-336-257A-6	Sequence 6, Appli
23	35	100.0	524	6 5386025-4	Patent No. 5386025
24	35	100.0	524	6 5386025-4	Patent No. 5386025
25	35	100.0	554	3 US-08-895-590-9	Sequence 9, Appli
26	35	100.0	702	4 US-09-917-254-94	Sequence 94, Appl
27	35	100.0	1088	3 US-09-082-059-2	Sequence 2, Appli

28	35	100.0	1250	3 US-08-938-291A-9	Sequence 9, Appli
29	35	100.0	1250	4 US-09-589-619-9	Sequence 9, Appli
30	35	100.0	1495	4 US-09-697-898-3	Sequence 3, Appli
31	35	100.0	1512	4 US-09-697-898-2	Sequence 2, Appli
32	35	100.0	4377	4 US-09-949-016-6978	Sequence 6978, Ap
33	29	82.9	10	3 US-09-460-384-22	Sequence 22, Appl
34	29	82.9	12	2 US-08-556-597-138	Sequence 138, Appl
35	29	82.9	12	4 US-10-118-575A-17	Sequence 17, Appl
36	29	82.9	15	3 US-08-602-999A-382	Sequence 382, App
37	29	82.9	15	4 US-09-500-124-382	Sequence 382, App
38	29	82.9	16	6 5378805-5	Patent No. 5378805
39	29	82.9	16	6 5378805-5	Patent No. 5378805
40	29	82.9	17	2 US-08-982-597A-16	Sequence 16, Appl
41	29	82.9	17	3 US-09-136-218-16	Sequence 16, Appl
42	29	82.9	17	3 US-08-602-999A-416	Sequence 416, App
43	29	82.9	17	4 US-09-500-124-416	Sequence 416, App
44	29	82.9	18	3 US-08-602-999A-303	Sequence 303, App
45	29	82.9	18	3 US-08-602-999A-371	Sequence 371, App

ALIGNMENTS

RESULT 1
US-07-745-206A-28
; Sequence 28, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-745-206A-28

Query Match 100.0%; Score 35; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5

Db 40 PPPHG 44

RESULT 2
US-08-311-363-28
; Sequence 28, Application US/083111363
; Patent No. 5876958
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,363
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-51506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-311-363-28

Query Match 100.0%; Score 35; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|
|
|
|
Db 40 PPPHG 44

RESULT 3
US-09-489-039A-10032
; Sequence 10032, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10032
; LENGTH: 98

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10032

Query Match 100.0%; Score 35; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|
|
|
|
Db 28 PPPHG 32

RESULT 4
US-09-489-039A-12696
; Sequence 12696, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12696
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12696

Query Match 100.0%; Score 35; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|
|
|
|
Db 10 PPPHG 14

RESULT 5
US-09-270-767-32232
; Sequence 32232, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32232
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32232

Query Match 100.0%; Score 35; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|
|
|
|
Db 36 PPPHG 40

RESULT 6
US-09-270-767-47449
; Sequence 47449, Application US/09270767
; Patent No. 6703491


```

; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47449
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47449

Query Match      100.0%; Score 35; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      36 PPPHG 40

RESULT 7
US-08-683-262B-47
; Sequence 47, Application US/08683262B
; Patent No. 5929220
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,262B
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-262B-47

Query Match      100.0%; Score 35; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      68 PPPHG 72

RESULT 8
US-08-683-262B-47
; Sequence 47, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; Li, Jisu
; Wands, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 00786/287003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-361-707-47

Query Match      100.0%; Score 35; DB 3; Length 174;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      68 PPPHG 72

RESULT 9
US-09-640-211A-1165
; Sequence 1165, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1165
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-640-211A-1165
```

```
Query Match          100.0%; Score 35; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      196 PPPHG 200

RESULT 10
PCT-US93-00227-4
; Sequence 4, Application PC/TUS9300227
; GENERAL INFORMATION:
; APPLICANT: Furneaux,, Henry M.
; APPLICANT: Posner, Jerome B.
; TITLE OF INVENTION: ANTIGEN RECOGNIZED BY PATIENTS WITH
; TITLE OF INVENTION: ANTIBODY ASSOCIATED LAMBERT-EATON MYASTHENIC SYNDROME
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00227
; FILING DATE: 19930111
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39874
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..240
;
PCT-US93-00227-4

Query Match          100.0%; Score 35; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      222 PPPHG 226

RESULT 11
US-09-270-767-43608
```

```
; Sequence 43608, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43608
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43608

Query Match          100.0%; Score 35; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      224 PPPHG 228

RESULT 12
US-09-252-991A-32545
; Sequence 32545, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32545
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32545

Query Match          100.0%; Score 35; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      90 PPPHG 94

RESULT 13
US-09-270-767-45004
; Sequence 45004, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45004
; LENGTH: 332
; TYPE: PRT
```

```
; ORGANISM: Drosophila melanogaster
US-09-270-767-45004

Query Match      100.0%; Score 35; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      184 PPPHG 188

RESULT 14
US-08-894-139-4
; Sequence 4, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-894-139-4

Query Match      100.0%; Score 35; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      20 PPPHG 24

RESULT 15
US-09-949-016-11268
; Sequence 11268, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11268
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11268

Query Match      100.0%; Score 35; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      308 PPPHG 312

RESULT 16
US-09-949-016-6701
; Sequence 6701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6701
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6701

Query Match      100.0%; Score 35; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      29 PPPHG 33

RESULT 17
US-09-248-796A-18484
; Sequence 18484, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; WITH HUMAN DISEASE, METHODS OF DETECTION AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

```

; SEQ ID NO 18484
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18484

Query Match      100.0%; Score 35; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      343 PPHG 347

RESULT 18
US-08-149-097D-38
; Sequence 38, Application US/08149097D
; Patent No. 5874236
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,097D
; FILING DATE: 05-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US92/06903
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,231
; FILING DATE: 13-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-55038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 479 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; FEATURE:
; OTHER INFORMATION: /standard_name= "Beta1-1"
US-08-149-097D-38

Query Match      100.0%; Score 35; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      249 PPHG 253

RESULT 19
US-10-029-180-106
; Sequence 106, Application US/10029180
; Patent No. 6806082
; GENERAL INFORMATION:
; APPLICANT: Cali, Brian M.
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin T.
; APPLICANT: Milna, G. Todd
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeffry C.
; APPLICANT: Trueheart, Josh
; APPLICANT: Zhang, Lixin
; TITLE OF INVENTION: No. 6806082el Regulators of Fungal Gene Expression
; FILE REFERENCE: MIC-004
; CURRENT APPLICATION NUMBER: US/10/029,180
; CURRENT FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: US 60/257,431
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 106
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fungal gene
US-10-029-180-106

Query Match      100.0%; Score 35; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      278 PPHG 282

RESULT 20
US-08-455-543A-42
```

; Sequence 42, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,543A
; FILING DATE: May 31, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/223,305
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-52517
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-455-543A-42

Query Match 100.0%; Score 35; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPHG 5

Db 249 PPHG 253
RESULT 21
US-08-223-305C-42
; Sequence 42, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/223,305C
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-223-305C-42

Query Match 100.0%; Score 35; DB 2; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 249 PPPHG 253

RESULT 22

US-08-336-257A-6
; Sequence 6, Application US/08336257A
; Patent No. 5726035
; GENERAL INFORMATION:
; APPLICANT: Jay, Scott D
; APPLICANT: Ellis, Steven B.
; APPLICANT: Harpold, Michael M.
; APPLICANT: Campbell, Kevin P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,257A
; FILING DATE: 07-NOV-1994
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 54898
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-336-257A-6

Query Match 100.0%; Score 35; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 249 PPPHG 253

RESULT 23

5386025-4
; Patent No. 5386025
; APPLICANT: JAY, SCOTT D.; ELLIS, STEVEN B.; HARPOLD, MICHAEL
; M.; CAMPBELL, KEVIN P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/482,384
; FILING DATE: 20-FEB-1990
; SEQ ID NO:4:
; LENGTH: 524

5386025-4

Query Match 100.0%; Score 35; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 249 PPPHG 253

RESULT 24

5386025-4
; Patent No. 5386025
; APPLICANT: JAY, SCOTT D.; ELLIS, STEVEN B.; HARPOLD, MICHAEL
; M.; CAMPBELL, KEVIN P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/482,384
; FILING DATE: 20-FEB-1990
; SEQ ID NO:4:
; LENGTH: 524
5386025-4

Query Match 100.0%; Score 35; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 249 PPPHG 253

RESULT 25

US-08-895-590-9
; Sequence 9, Application US/088955590
; Patent No. 6207410
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,590
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,888
; FILING DATE: 19-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-263
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:


```

; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-895-590-9

Query Match      100.0%; Score 35; DB 3; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      250 PPPHG 254

RESULT 26
US-09-917-254-94
; Sequence 94, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-94

Query Match      100.0%; Score 35; DB 4; Length 702;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      228 PPPHG 232

RESULT 27
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devarajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

Query Match      100.0%; Score 35; DB 3; Length 1088;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
```

```

Db      627 PPPHG 631

RESULT 28
US-08-938-291A-9
; Sequence 9, Application US/08938291A
; Patent No. 6117673
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: RDGB PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,291A
; FILING DATE: September 26, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,337
; FILING DATE: October 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 228/172
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-938-291A-9

Query Match      100.0%; Score 35; DB 3; Length 1250;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      435 PPPHG 439

RESULT 29
US-09-589-619-9
; Sequence 9, Application US/09589619
; Patent No. 6576442
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: RDGB PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/589,619
FILING DATE: 07-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/938,291
FILING DATE: September 26, 1997
APPLICATION NUMBER: 60/027,337
FILING DATE: October 11, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 228/172
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1250 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-589-619-9
Query Match 100.0%; Score 35; DB 4; Length 1250;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db 435 PPPHG 439
RESULT 30
US-09-697-898-3
; Sequence 3, Application US/09697898
; Patent No. 6818427
; GENERAL INFORMATION:
; APPLICANT: Palombella, Vito J.
; APPLICANT: Liao, Sha-Mei
; TITLE OF INVENTION: MEKK1 Molecules and Uses Thereof
; FILE REFERENCE: 103576.144
; CURRENT APPLICATION NUMBER: US/09/697,898
; CURRENT FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1495
; TYPE: PRT
; ORGANISM: partial human MEKK1 protein
US-09-697-898-3
Query Match 100.0%; Score 35; DB 4; Length 1495;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5

Db 92 PPPHG 96
RESULT 31
US-09-697-898-2
; Sequence 2, Application US/09697898
; Patent No. 6818427
; GENERAL INFORMATION:
; APPLICANT: Palombella, Vito J.
; APPLICANT: Liao, Sha-Mei
; TITLE OF INVENTION: MEKK1 Molecules and Uses Thereof
; FILE REFERENCE: 103576.144
; CURRENT APPLICATION NUMBER: US/09/697,898
; CURRENT FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1512
; TYPE: PRT
; ORGANISM: Human
US-09-697-898-2
Query Match 100.0%; Score 35; DB 4; Length 1512;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db 111 PPPHG 115
RESULT 32
US-09-949-016-6978
; Sequence 6978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6978
; LENGTH: 4377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6978
Query Match 100.0%; Score 35; DB 4; Length 4377;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db 1033 PPPHG 1037
RESULT 33
US-09-460-384-22
; Sequence 22, Application US/09460384
; Patent No. 6337316
; GENERAL INFORMATION:
; APPLICANT: EL TAYAR, Nabil
; APPLICANT: BLECHNER, Steven

APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: QUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 382:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-382
Query Match 82.9%; Score 29; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPH 4
Db 9 PPPH 12
RESULT 37
US-09-500-124-382
Sequence 382, Application US/09500124
Patent No. 6432920
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: QUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,124
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,999
FILING DATE: 16-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 382:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-500-124-382
Query Match 82.9%; Score 29; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPH 4
Db 9 PPPH 12
RESULT 38
5378805-5
Patent No. 5378805
APPLICANT: Lai, Renu B.
TITLE OF INVENTION: IMMUNOREACTIVE HTLV-I/II ENV AND POL PEPTIDES
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/574,352
FILING DATE: 29-AUG-1990
SEQ ID NO: 5:
LENGTH: 16
5378805-5
Query Match 82.9%; Score 29; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPH 4
Db 12 PPPH 15
RESULT 39
5378805-5
Patent No. 5378805
APPLICANT: Lai, Renu B.
TITLE OF INVENTION: IMMUNOREACTIVE HTLV-I/II ENV AND POL PEPTIDES
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/574,352
FILING DATE: 29-AUG-1990
SEQ ID NO: 5:
LENGTH: 16
5378805-5
Query Match 82.9%; Score 29; DB 6; Length 16;

Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPH 4
 ||||
Db 12 PPPH 15

RESULT 40
US-08-982-597A-16
; Sequence 16, Application US/08982597A
; Patent No. 5932693
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staatz, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 5932693th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,597A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,542
; FILING DATE: 10-DEC-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-3002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-3117
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-982-597A-16

Query Match 82.9%; Score 29; DB 2; Length 17;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
 |||.|
Db 12 PPYPG 16

Search completed: June 15, 2005, 14:24:21
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-10

Perfect score: 35

Sequence: 1 PPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	5	14	US-10-074-225A-10
2	35	100.0	40	16	US-10-425-115-307567
3	35	100.0	44	16	US-10-437-963-139202
4	35	100.0	48	16	US-10-425-115-340254
5	35	100.0	51	16	US-10-425-115-199641
6	35	100.0	53	16	US-10-425-115-242501
7	35	100.0	55	9	US-09-864-761-34953
8	35	100.0	55	15	US-10-424-599-170955
9	35	100.0	56	16	US-10-425-115-199279
10	35	100.0	58	16	US-10-437-963-152655
11	35	100.0	58	16	US-10-425-115-323782

12	35	100.0	59	16	US-10-425-115-261805	Sequence 261805,
13	35	100.0	60	16	US-10-425-115-352770	Sequence 352770,
14	35	100.0	61	9	US-09-864-761-46134	Sequence 46134, A
15	35	100.0	61	16	US-10-425-115-239349	Sequence 239349,
16	35	100.0	62	16	US-10-425-115-344565	Sequence 344565,
17	35	100.0	65	15	US-10-424-599-206662	Sequence 206662,
18	35	100.0	65	16	US-10-425-115-316672	Sequence 316672,
19	35	100.0	68	16	US-10-437-963-130521	Sequence 130521,
20	35	100.0	70	16	US-10-437-963-150972	Sequence 150972,
21	35	100.0	72	15	US-10-424-599-181807	Sequence 181807,
22	35	100.0	72	16	US-10-425-115-185365	Sequence 185365,
23	35	100.0	73	15	US-10-424-599-187344	Sequence 187344,
24	35	100.0	75	16	US-10-767-701-51220	Sequence 51220, A
25	35	100.0	75	16	US-10-425-115-307219	Sequence 307219,
26	35	100.0	76	16	US-10-425-115-258243	Sequence 258243,
27	35	100.0	77	16	US-10-437-963-158099	Sequence 158099,
28	35	100.0	79	16	US-10-437-963-132911	Sequence 132911,
29	35	100.0	80	15	US-10-424-599-220866	Sequence 220866,
30	35	100.0	80	16	US-10-425-115-249543	Sequence 249543,
31	35	100.0	80	16	US-10-425-115-323576	Sequence 323576,
32	35	100.0	81	16	US-10-437-963-109589	Sequence 109589,
33	35	100.0	81	16	US-10-425-115-233659	Sequence 233659,
34	35	100.0	82	16	US-10-425-115-271683	Sequence 271683,
35	35	100.0	84	15	US-10-424-599-163493	Sequence 163493,
36	35	100.0	84	15	US-10-424-599-275367	Sequence 275367,
37	35	100.0	84	16	US-10-437-963-133341	Sequence 133341,
38	35	100.0	86	16	US-10-425-115-204320	Sequence 204320,
39	35	100.0	88	15	US-10-424-599-143090	Sequence 143090,
40	35	100.0	88	15	US-10-424-599-260730	Sequence 260730,
41	35	100.0	89	15	US-10-282-122A-62241	Sequence 62241, A
42	35	100.0	89	16	US-10-425-115-232815	Sequence 232815,
43	35	100.0	90	16	US-10-437-963-123295	Sequence 123295,
44	35	100.0	91	15	US-10-424-599-151231	Sequence 151231,
45	35	100.0	91	15	US-10-424-599-252891	Sequence 252891,

ALIGNMENTS

RESULT 1
US-10-074-225A-10
; Sequence 10, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-10

Query Match 100.0%; Score 35; DB 14; Length 5;
Best local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 1 PPPHG 5

RESULT 2
US-10-425-115-307567
; Sequence 307567, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 307567
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_43570C.1.pep
US-10-425-115-307567

Query Match 100.0%; Score 35; DB 16; Length 40;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 14 PPPHG 18

RESULT 3
US-10-437-963-139202
; Sequence 139202, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 139202
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(44)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40517C.1.pep
US-10-437-963-139202

Query Match 100.0%; Score 35; DB 16; Length 44;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 30 PPPHG 34

RESULT 4
US-10-425-115-340254
; Sequence 340254, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 340254
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_73481C.1.pep
US-10-425-115-340254

Query Match 100.0%; Score 35; DB 16; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 43 PPPHG 47

RESULT 5
US-10-425-115-199641
; Sequence 199641, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199641
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_113645C.1.pep
US-10-425-115-199641

Query Match 100.0%; Score 35; DB 16; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 31 PPPHG 35

RESULT 6
US-10-425-115-242501
; Sequence 242501, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 242501
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_152747C.1.pep
US-10-425-115-242501

Query Match 100.0%; Score 35; DB 16; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 31 PPPHG 35

RESULT 7
US-09-864-761-34953
; Sequence 34953, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34953
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004900.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EST HUMAN HIT: AW953375.1, EVALUE 6.00e-29
; OTHER INFORMATION: SWISSPROT HIT: P16793, EVALUE 5.20e-01
US-09-864-761-34953

Query Match 100.0%; Score 35; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 1 PPPHG 5

RESULT 8
US-10-424-599-170955
; Sequence 170955, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170955
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125388C.1.pap
US-10-424-599-170955

Query Match 100.0%; Score 35; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 33 PPPHG 37

RESULT 9
US-10-425-115-199279
; Sequence 199279, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199279
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_113318C.1.pep
US-10-425-115-199279

Query Match      100.0%; Score 35; DB 16; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      1 PPPHG 5

RESULT 10
US-10-437-963-152655
; Sequence 152655, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152655
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_52685C.1.pep
US-10-437-963-152655

Query Match      100.0%; Score 35; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      21 PPPHG 25

RESULT 11
US-10-425-115-323782
; Sequence 323782, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
```

```

; SEQ ID NO 323782
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_5835C.1.pep
US-10-425-115-323782

Query Match      100.0%; Score 35; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      11 PPPHG 15

RESULT 12
US-10-425-115-261805
; Sequence 261805, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261805
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170379C.1.pep
US-10-425-115-261805

Query Match      100.0%; Score 35; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
Db      49 PPPHG 53

RESULT 13
US-10-425-115-352770
; Sequence 352770, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 352770
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84894C.1.pep
US-10-425-115-352770

Query Match      100.0%; Score 35; DB 16; Length 60;
```

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
 |||||
Db 31 PPPHG 35

RESULT 14
US-09-864-761-46134
; Sequence 46134, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00658
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46134
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011625.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
US-09-864-761-46134

Query Match 100.0%; Score 35; DB 9; Length 61;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
 |||||
Db 18 PPPHG 22

RESULT 15
US-10-425-115-239349
; Sequence 239349, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 239349
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149873C.1.pep
US-10-425-115-239349

Query Match 100.0%; Score 35; DB 16; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
 |||||
Db 2 PPPHG 6

RESULT 16
US-10-425-115-344565
; Sequence 344565, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 344565
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77405C.1.pep
US-10-425-115-344565

Query Match 100.0%; Score 35; DB 16; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
 |||||
Db 27 PPPHG 31

RESULT 17

US-10-424-599-206662
; Sequence 206662, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206662
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28643C.1.1.pep
US-10-424-599-206662

Query Match 100.0%; Score 35; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 55 PPPHG 59

RESULT 18
US-10-425-115-316672
; Sequence 316672, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 316672
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51874C.1.1.pep
US-10-425-115-316672

Query Match 100.0%; Score 35; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 25 PPPHG 29

RESULT 19
US-10-437-963-130521
; Sequence 130521, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130521
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32676C.1.1.pep
US-10-437-963-130521

Query Match 100.0%; Score 35; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 48 PPPHG 52

RESULT 20
US-10-437-963-150972
; Sequence 150972, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 150972
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_51159C.1.1.pep
US-10-437-963-150972

Query Match 100.0%; Score 35; DB 16; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 13 PPPHG 17

RESULT 21
US-10-424-599-181807
; Sequence 181807, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement


```

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181807
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135185C.1.pep
US-10-424-599-181807

Query Match      100.0%; Score 35; DB 15; Length 72;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      7 PPPHG 11

RESULT 22
US-10-425-115-185365
; Sequence 185365, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 185365
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(72)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_100640C.1.pep
US-10-425-115-185365

Query Match      100.0%; Score 35; DB 16; Length 72;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      42 PPPHG 46

RESULT 23
US-10-424-599-187344
; Sequence 187344, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187344
```

```

; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140183C.1.pep
US-10-424-599-187344

Query Match      100.0%; Score 35; DB 15; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      69 PPPHG 73

RESULT 24
US-10-767-701-51220
; Sequence 51220, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 51220
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3481-060-Q6-K1-D12.pep
US-10-767-701-51220

Query Match      100.0%; Score 35; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPPHG 5
      |||||
Db      8 PPPHG 12

RESULT 25
US-10-425-115-307219
; Sequence 307219, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 307219
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(75)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_43257C.1.pep
US-10-425-115-307219
```

Query Match 100.0%; Score 35; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 5 PPPHG 9

RESULT 26
US-10-425-115-258243
; Sequence 158099, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 258243
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167102C.1.pep
US-10-425-115-258243

Query Match 100.0%; Score 35; DB 16; Length 76;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 29 PPPHG 33

RESULT 27
US-10-437-963-158099
; Sequence 158099, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 158099
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57606C.1.pep
US-10-437-963-158099

Query Match 100.0%; Score 35; DB 16; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 16 PPPHG 20

RESULT 28
US-10-437-963-132911
; Sequence 132911, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 132911
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_34833C.1.pep
US-10-437-963-132911

Query Match 100.0%; Score 35; DB 16; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 37 PPPHG 41

RESULT 29
US-10-424-599-220866
; Sequence 220866, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220866
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(80)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41472C.1.pep
US-10-424-599-220866

Query Match 100.0%; Score 35; DB 15; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5


```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 271683
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_179368C.1.pep
US-10-425-115-271683

Query Match          100.0%; Score 35; DB 16; Length 82;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      44 PPPHG 48

RESULT 35
US-10-424-599-163493
; Sequence 163493, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163493
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118653C.1.pep
US-10-424-599-163493

Query Match          100.0%; Score 35; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      38 PPPHG 42

RESULT 36
US-10-424-599-275367
; Sequence 275367, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
```

```
; SEQ ID NO 275367
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90677C.1.pep
US-10-424-599-275367

Query Match          100.0%; Score 35; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      69 PPPHG 73

RESULT 37
US-10-437-963-133341
; Sequence 133341, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133341
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35220C.1.pep
US-10-437-963-133341

Query Match          100.0%; Score 35; DB 16; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      7 PPPHG 11

RESULT 38
US-10-425-115-204320
; Sequence 204320, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 204320
```

; LENGTH: 86
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_117932C.1.pep
US-10-425-115-204320

Query Match 100.0%; Score 35; DB 16; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 82 PPPHG 86

RESULT 39
US-10-424-599-143090
; Sequence 143090, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143090
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100223C.1.pep
US-10-424-599-143090

Query Match 100.0%; Score 35; DB 15; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 7 PPPHG 11

RESULT 40
US-10-424-599-260730
; Sequence 260730, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260730
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77461C.1.pep
US-10-424-599-260730

Query Match 100.0%; Score 35; DB 15; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PPPHG 5
Db 61 PPPHG 65

Search completed: June 15, 2005, 14:50:21
Job time : 81.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-10
Perfect score: 35
Sequence: 1 PPPHG 5
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	111	2 C72714	hypothetical prote
2	35	100.0	147	2 S37485	gene msgl protein
3	35	100.0	168	2 C87235	conserved hypothet
4	35	100.0	169	2 D70906	hypothetical prote
5	35	100.0	170	2 A48013	proline-rich prote
6	35	100.0	173	2 T51469	glycine/proline-ri
7	35	100.0	173	2 T19341	hypothetical prote
8	35	100.0	175	2 T46665	hypothetical prote
9	35	100.0	176	2 A86441	hypothetical prote
10	35	100.0	206	1 PIRT3	acidic proline-ric
11	35	100.0	215	2 F72641	hypothetical prote
12	35	100.0	224	2 B87440	transcription regu
13	35	100.0	257	2 T10586	small nuclear ribo
14	35	100.0	278	1 EDBE63	immediate-early pr
15	35	100.0	301	2 T35805	hypothetical prote
16	35	100.0	304	2 A55852	membrane-associate
17	35	100.0	316	2 S16681	homeotic protein -
18	35	100.0	335	2 I48338	E2F-5 - mouse
19	35	100.0	366	2 T26449	hypothetical prote
20	35	100.0	368	2 G84769	hypothetical prote
21	35	100.0	391	2 G87723	protein R06A10.4 [
22	35	100.0	422	2 T07967	mandelonitrile lya
23	35	100.0	429	2 T32632	hypothetical prote
24	35	100.0	434	2 S19010	homeotic protein P
25	35	100.0	454	2 F81046	hypothetical prote
26	35	100.0	459	2 B69458	hypothetical prote
27	35	100.0	465	2 G02738	sulfate adenyltr
28	35	100.0	475	2 A44461	FREAC-4 - human
29	35	100.0	486	2 T20482	voltage-dependent hypothetical prote

30	35	100.0	509	2 JC6203	SP8 binding protei
31	35	100.0	516	2 T49422	RAD57 related prot
32	35	100.0	518	2 S61920	B-alpha pheromone-
33	35	100.0	522	2 I65767	L-type voltage-gat
34	35	100.0	524	2 A41347	calcium channel pr
35	35	100.0	529	2 T00677	hypothetical prote
36	35	100.0	597	2 G84825	probable CCCH-type
37	35	100.0	614	2 E86194	hypothetical prote
38	35	100.0	627	2 S48424	hypothetical prote
39	35	100.0	629	2 C87048	probable ABC trans
40	35	100.0	669	2 A49585	Na+ channel protei
41	35	100.0	830	2 T36683	probable integral
42	35	100.0	960	2 A41638	chitin synthase (E
43	35	100.0	968	2 F70746	probable mmpL2 pro
44	35	100.0	968	2 T00353	hypothetical prote
45	35	100.0	1048	2 T31425	C-terminal domain-

ALIGNMENTS

RESULT 1

C72714
hypothetical protein APE1130 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: C72714
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72714
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-111 <KAW>
A;Cross-references: UNIPROT:Q9YCY2; DDBJ:AP000060; NID:g5104188; PIDN:BAA80115.1; PID:d1
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1130

Query Match 100.0%; Score 35; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 82 PPPHG 86

RESULT 2

S37485
gene msgl protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: I48669; S37485
R;Tronik-Le Roux, D.; Senorale-Pose, M.; Rougeon, F.
Gene 142, 175-182, 1994
A;Title: Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice
A;Reference number: I48669; MUID:94252564; PMID:8194749
A;Accession: I48669
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-147 <RES>
A;Cross-references: UNIPROT:Q61900; EMBL:X71629; NID:g406256; PIDN:CAA50636.1; PID:g40625
C;Genetics:
A;Gene: msgl
C;Superfamily: proline-rich peptide P-B

Query Match 100.0%; Score 35; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
| | | | |
Db 39 PPPHG 43

RESULT 3
C87235
conserved hypothetical protein ML2605 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87235
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: C87235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-168 <STO>
A;Cross-references: UNIPROT:O06065; GB:AL450380; NID:g13093836; PIDN:CAC32137.1; GSPDB:G
C;Genetics:
A;Gene: ML2605
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0185

Query Match 100.0%; Score 35; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
| | | | |
Db 131 PPPHG 135

RESULT 4
D70906
hypothetical protein Rv0185 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: D70906
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70906
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-169 <COL>
A;Cross-references: UNIPROT:O07429; GB:Z97050; GB:AL123456; NID:g3256008; PIDN:CAB09736.
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0185
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0185

Query Match 100.0%; Score 35; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
| | | | |
Db 132 PPPHG 136

RESULT 5
A48013
proline-rich proteoglycan 1 precursor, parotid - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 09-Jul-2004

C;Accession: A48013
R;Castle, A.M.; Castle, J.D.
J. Biol. Chem. 268, 20490-20496, 1993
A;Title: Novel secretory proline-rich proteoglycans from rat parotid. Cloning and charact
A;Reference number: A48013; MUID:93388626; PMID:8376404
A;Accession: A48013
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-170 <CAS>
A;Cross-references: UNIPROT:Q07610; GB:LI17317; NID:g310197; PIDN:AAA03073.1; PID:g310198
C;Superfamily: proline-rich protein
C;Keywords: extracellular protein; glycoprotein; tandem repeat

Query Match 100.0%; Score 35; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
| | | | |
Db 80 PPPHG 84

RESULT 6
T51469
glycine/proline-rich protein - Arabidopsis thaliana
N;Alternate names: protein K10A8_130
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51469
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mewe
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51469
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <SAT>
A;Cross-references: UNIPROT:Q9LF59; EMBL:AL391151
A;Experimental source: cultivar Columbia; BAC clone K10A8
C;Genetics:
A;Map position: 5
A;Introns: 97/1
A;Note: K10A8_130

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
| | | | |
Db 49 PPPHG 53

RESULT 7
T19341
hypothetical protein Cl6D6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19341
R;Gardner, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19111
A;Accession: T19341
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-173 <WIL>
A;Cross-references: UNIPROT:O62063; EMBL:Z81472; PIDN:CAB03889.1; GSPDB:GN00028; CESP:C16
A;Experimental source: clone Cl6D6
C;Genetics:
A;Gene: CESP:Cl6D6.1
A;Map position: X
A;Introns: 42/1

Query Match 100.0%; Score 35; DB 2; Length 173;

Best Local Similarity 100.0%; Pred. No. 44; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5
Db	72 PPPHG 76
RESULT 8 T46665 hypothetical protein [imported] - Rhodococcus corallinus (fragment) C;Species: Rhodococcus corallinus C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004 C;Accession: T46665 R;Shao, Z.Q.; Sefens, W.; Mulbry, W.; Behki, R.M. J. Bacteriol. 177, 5748-5755, 1995 A;Title: Cloning and expression of the s-triazine hydrolase gene (trza) from Rhodococcus erbicide atrazine. A;Reference number: Z23125; MUID:96011356; PMID:7592318 A;Accession: T46665 A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: DNA A;Residues: 1-175 <SHA> A;Cross-references: UNIPROT:Q52724; EMBL:L16534; NID:g294669; PIDN:AAA90930.1; PID:g294669 A;Experimental source: tissue lib NRRL 15444B	
Query Match 100.0%; Score 35; DB 2; Length 175; Best Local Similarity 100.0%; Pred. No. 45; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5
Db	129 PPPHG 133
RESULT 9 A86441 hypothetical protein F5M6.24 - Arabidopsis thaliana C;Species: Arabidopsis thaliana (mouse-ear cress) C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004 C;Accession: A86441 R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H. A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W. A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis. A;Reference number: A86141; MUID:21016719; PMID:11130712 A;Accession: A86441 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-176 <STO> A;Cross-references: UNIPROT:Q9C4Z8; GB:AE005172; NID:g11136740; PIDN:AAG31321.1; GSPDB:G C;Genetics: A;Map position: 1	
Query Match 100.0%; Score 35; DB 2; Length 176; Best Local Similarity 100.0%; Pred. No. 45; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5
Db	50 PPPHG 54
RESULT 10 PIRT3 acidic proline-rich protein precursor - rat N;Alternate names: PRP	

C;Species: Rattus norvegicus (Norway rat) C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004 C;Accession: A03296 R;Ziemer, M.A.; Swain, W.F.; Rutter, W.J.; Clements, S.; Ann, D.K.; Carlson, D.M. J. Biol. Chem. 259, 10475-10480, 1984 A;Title: Nucleotide sequence analysis of a proline-rich protein cDNA and peptide homolog; A;Reference number: A03296; MUID:84289443; PMID:6547951 A;Accession: A03296 A;Molecule type: mRNA A;Residues: 1-206 <ZIE> A;Cross-references: UNIPROT:P04474; GB:K02247; NID:g206395; PIDN:AAA41949.1; PID:g206396 C;Comment: This protein contains six 18- to 19-residue repeats. C;Comment: This protein may protect teeth by binding to tannins. C;Superfamily: proline-rich protein C;Keywords: duplication; parotid gland; saliva; tandem repeat F;1-13/Domain: signal sequence #status predicted <SIG> F;14-206/Product: acidic proline-rich protein #status predicted <MAT> F;80-189/Region: 18-residue repeats	
Query Match 100.0%; Score 35; DB 1; Length 206; Best Local Similarity 100.0%; Pred. No. 53; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5
Db	80 PPPHG 84
RESULT 11 F72641 hypothetical protein APE0566 - Aeropyrum pernix (strain K1) C;Species: Aeropyrum pernix C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004 C;Accession: F72641 R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki DNA Res. 6, 83-101, 1999 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr A;Reference number: A72450; MUID:99310339; PMID:10382966 A;Accession: F72641 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-215 <KAN> A;Cross-references: UNIPROT:Q9YEL2; DDBJ:AP000060; NID:g5104188; PIDN:BAA79534.1; PID:g51 A;Experimental source: strain K1 C;Genetics: A;Gene: APE0566 C;Superfamily: proline-rich protein	
Query Match 100.0%; Score 35; DB 2; Length 215; Best Local Similarity 100.0%; Pred. No. 55; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5
Db	57 PPPHG 61
RESULT 12 B87440 transcription regulator, TetR family [imported] - Caulobacter crescentus C;Species: Caulobacter crescentus C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 C;Accession: B87440 R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. .B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001 A;Title: Complete Genome Sequence of Caulobacter crescentus. A;Reference number: AB7249; MUID:21173698; PMID:11259647 A;Accession: B87440 A;Status: preliminary A;Molecule type: DNA	

A;Residues: 1-224 <STO>
A;Cross-references: UNIPROT:Q9A825; GB:AE005673; NID:g13422922; PIDN:AAK23518.1; GSPDB:C
C;Genetics:
A;Gene: CC1539

Query Match 100.0%; Score 35; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 219 PPPHG 223

RESULT 13
T10586
small nuclear ribonucleoprotein-associated protein homolog F9F13.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10586
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16991
A;Accession: T10586
A;Molecule type: DNA
A;Residues: 1-257 <BEV>
A;Cross-references: UNIPROT:Q9SUN5; EMBL:AL080253; GSPDB:GN00062; ATSP:F9F13.90
A;Experimental source: cultivar Columbia; BAC clone F9F13
C;Genetics:
A;Gene: ATSP:F9F13.90
A;Map position: 4
C;Superfamily: proline-rich protein

Query Match 100.0%; Score 35; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 221 PPPHG 225

RESULT 14
FDBE63
immediate-early protein IE68 - human herpesvirus 3
C;Species: human herpesvirus 3, varicella-zoster virus
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C;Accession: B27345; I27345
R;Davison, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A;Title: The complete DNA sequence of varicella-zoster virus.
A;Reference number: A27345; MUID:86306657; PMID:3018124
A;Accession: B27345
A;Molecule type: DNA
A;Residues: 1-278 <DAV>
A;Cross-references: UNIPROT:P09255; EMBL:X04370; NID:g59989; PIDN:CAA27946.1; PID:g60052
C;Genetics:
A;Gene: 63; 70
C;Superfamily: herpesvirus immediate-early protein IE68
C;Keywords: early protein

Query Match 100.0%; Score 35; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 273 PPPHG 277

RESULT 15
T35805
hypothetical protein SC8D9.01c SC8D9.01c - Streptomyces coelicolor (fragment)

C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 12-Jul-2004
C;Accession: T35805
R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A;Reference number: Z21589
A;Accession: T35805
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-301 <MUR>
A;Cross-references: UNIPROT:Q8CJR2; EMBL:AL035569; PIDN:CAB37565.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC8D9.01c
C;Superfamily: tropinesterase

Query Match 100.0%; Score 35; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 152 PPPHG 156

RESULT 16
A55852
membrane-associated protein A - Synechococcus sp.
C;Species: Synechococcus sp.
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
C;Accession: A55852
R;Webb, R.; Troyan, T.; Sherman, D.; Sherman, L.A.
J. Bacteriol. 176, 4906-4913, 1994
A;Title: MapA, an iron-regulated, cytoplasmic membrane protein in the cyanobacterium Syne
A;Reference number: A55852; MUID:94327456; PMID:8051004
A;Accession: A55852
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-304 <WEB>
A;Cross-references: UNIPROT:Q55277; GB:L01621; NID:g484231; PIDN:AAA59057.1; PID:g484232
C;Genetics:
A;Gene: mapA

Query Match 100.0%; Score 35; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 169 PPPHG 173

RESULT 17
S16681
homeotic protein - human
C;Species: Homo sapiens (man)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2004
C;Accession: S16681
R;Deguchi, Y.; Kehrl, J.H.
Nucleic Acids Res. 19, 3742, 1991
A;Title: Nucleotide sequence of a novel diverged human homeobox gene encodes a DNA bindi
A;Reference number: S16681; MUID:91305125; PMID:1677181
A;Accession: S16681
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-316 <DEG>
A;Cross-references: EMBL:X56537
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1991
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;73-127/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 35; DB 2; Length 316;

Best Local Similarity 100.0%; Pred. No. 81; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPHG 5
Db	270 PPHG 274
RESULT 18 I48338	
E2F-5 - mouse	
C;Species: Mus musculus (house mouse)	
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004	
C;Accession: I48338	
R;Buck, V.; Allen, K.E.; Sorensen, T.; Bybee, A.; Hijmans, E.M.; Voorhoeve, P.M.; Berna	
Oncogene 11, 31-38, 1995	
A;Title: Molecular and functional characterisation of E2F-5, a new member of the E2F fam	
A;Reference number: I48338; MUID:95349934; PMID:7542760	
A;Accession: I48338	
A;Status: preliminary; translated from GB/EMBL/DBJ	
A;Molecule type: mRNA	
A;Residues: 1-335 <RES>	
A;Cross-references: UNIPROT:Q61502; EMBL:X86925; NID:g806571; PIDN:CAA60508.1; PID:g8065	
C;Genetics:	
A;Gene: e2f-5	
Query Match 100.0%; Score 35; DB 2; Length 335; Best Local Similarity 100.0%; Pred. No. 86; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPHG 5
Db	20 PPHG 24
RESULT 19 T26449	
hypothetical protein Y113G7B.23 - Caenorhabditis elegans	
C;Species: Caenorhabditis elegans	
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004	
C;Accession: T26449	
R;Lennard, N.	
submitted to the EMBL Data Library, September 1999	
A;Reference number: Z20215	
A;Accession: T26449	
A;Status: preliminary; translated from GB/EMBL/DBJ	
A;Molecule type: DNA	
A;Residues: 1-366 <WIL>	
A;Cross-references: UNIPROT:Q9U2W6; EMBL:AL110477; NID:e1542121; PIDN:CAB54337.1; CESP:Y	
A;Experimental source: clone Y113G7B	
C;Genetics:	
A;Gene: CESP:Y113G7B.23	
A;Introns: 318/1	
Query Match 100.0%; Score 35; DB 2; Length 366; Best Local Similarity 100.0%; Pred. No. 94; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPHG 5
Db	330 PPHG 334
RESULT 20 G84769	
hypothetical protein At2g35530 [imported] - Arabidopsis thaliana	
C;Species: Arabidopsis thaliana (mouse-ear cress)	
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004	
C;Accession: G84769	
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;	
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.	
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J	
Nature 402, 761-768, 1999	

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.	
A;Reference number: A84420; MUID:20083487; PMID:10617197	
A;Accession: G84769	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-368 <STO>	
A;Cross-references: UNIPROT:O82288; GB:AE002093; NID:g3608135; PIDN:AAC36168.1; GSPDB:GN	
C;Genetics:	
A;Gene: At2g35530	
A;Map position: 2	
C;Superfamily: fava bean G-box-binding protein; fos/jun DNA-binding domain homology	
Query Match 100.0%; Score 35; DB 2; Length 368; Best Local Similarity 100.0%; Pred. No. 94; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPHG 5
Db	54 PPHG 58
RESULT 21 G87723	
protein R06A10.4 [imported] - Caenorhabditis elegans	
C;Species: Caenorhabditis elegans	
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004	
C;Accession: G87723	
R;anonymous, The C. elegans Sequencing Consortium.	
Science 282, 2012-2018, 1998	
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology	
A;Reference number: A75000; MUID:99069613; PMID:9851916	
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele	
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and	
A;Accession: G87723	
A;Status: preliminary	
A;Molecule type: DNA	
A;Residues: 1-391 <STO>	
A;Cross-references: UNIPROT:O44747; GB:chr_I; PIDN:AAB96730.1; PID:g2773215; GSPDB:GN0000	
C;Genetics:	
A;Gene: R06A10.4	
A;Map position: 1	
Query Match 100.0%; Score 35; DB 2; Length 391; Best Local Similarity 100.0%; Pred. No. 1e+02; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPHG 5
Db	40 PPHG 44
RESULT 22 T07967	
mandelonitrile lyase (EC 4.1.2.10) - flax	
N;Alternate names: hydroxynitrile lyase	
C;Species: Linum usitatissimum (flax)	
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004	
C;Accession: T07967	
R;Breithaupt, H.M.; Boenigk, W.; Pohl, M.; Kula, M.R.	
submitted to the EMBL Data Library, September 1997	
A;Description: Cloning and overexpression of (R)-hydroxynitrile lyase from Linum usitat	
A;Reference number: Z16253	
A;Accession: T07967	
A;Status: preliminary; translated from GB/EMBL/DBJ	
A;Molecule type: mRNA	
A;Residues: 1-422 <BRE>	
A;Cross-references: UNIPROT:O22574; EMBL:AF024588; NID:g2565287; PIDN:AAB81956.1; PID:g2:	
C;Genetics:	
A;Gene: HNL	
C;Function:	
A;Description: catalyzes reversible addition of HCN to aliphatic ketones and aldehydes	
C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology	
C;Keywords: aldehyde-lyase; carbon-carbon lyase	

Query Match 100.0%; Score 35; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 190 PPPHG 194

RESULT 23
T32632
hypothetical protein F15E6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32632
R;Miller, N.; Stellyes, L.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid F15E6.
A;Reference number: Z21202
A;Accession: T32632
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-429 <MIL>
A;Cross-references: UNIPROT:Q44497; EMBL:AF038614; PIDN:AAB92062.1; GSPDB:GN00022; CESP:
A;Experimental source: strain Bristol N2; clone F15E6
C;Genetics:
A;Gene: CESP:F15E6.2
A;Map position: 4
A;Introns: 196/3; 241/3; 338/1

Query Match 100.0%; Score 35; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 369 PPPHG 373

RESULT 24
S19010
homeotic protein PBX3a - human
N;Alternate names: homeobox protein; pre-B-cell leukemia transcription factor 3
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2004
C;Accession: S19010
R;Monica, K.; Galili, N.; Nourse, J.; Saltman, D.; Cleary, M.L.
Mol. Cell. Biol. 11, 6149-6157, 1991
A;Title: PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-on
A;Reference number: S19009; MUID:92049345; PMID:1682799
A;Accession: S19010
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-434 <MON>
A;Cross-references: UNIPROT:P40426; EMBL:X59841; NID:g35314; PIDN:CAA42502.1; PID:g35315
C;Genetics:
A;Gene: GDB:PBX3
A;Cross-references: GDB:125353; OMIM:176312
A;Map position: 9q33-9q34
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;236-295/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 35; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 29 PPPHG 33

RESULT 25
F81046
hypothetical protein NMB1759 - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: F81046
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: F81046
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-454 <TET>
A;Cross-references: UNIPROT:Q9JY39; GB:AE002525; GB:AE002098; NID:g7227004; PIDN:AAF4210
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1759
C;Superfamily: Neisseria meningitidis hypothetical protein NMB1759

Query Match 100.0%; Score 35; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 234 PPPHG 238

RESULT 26
B69458
sulfate adenylyltransferase (sat) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C;Accession: B69458
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: B69458
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-459 <KLE>
A;Cross-references: GB:AE000988; GB:AE000782; NID:g2689311; PIDN:AAB89581.1; PID:g264888
C;Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology

Query Match 100.0%; Score 35; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 10 PPPHG 14

RESULT 27
G02738
FREAC-4 - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Aug-2004
C;Accession: G02738
R;Enerback, S.
submitted to the EMBL Data Library, June 1996
A;Reference number: H01646
A;Accession: G02738
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA
A;Residues: 1-465 <ENE>
A;Cross-references: UNIPROT:Q16676; EMBL:U59832; NID:g13992238; PID:g13992239
C;Genetics:
A;Gene: FREAC-4
C;Superfamily: fork head DNA-binding domain homology
F;125-216/Domain: fork head DNA-binding domain homology <FHD>

Query Match 100.0%; Score 35; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 313 PPPHG 317

RESULT 28
A44461
voltage-dependent calcium channel beta-1 chain, M isoform - human
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C;Accession: A44461; I38001; S31919
R;Powers, P.A.; Liu, S.; Hogan, K.; Gregg, R.G.
J. Biol. Chem. 267, 22967-22972, 1992
A;Title: Skeletal muscle and brain isoforms of a beta-subunit of human voltage-dependent
A;Reference number: A44461; MUID:93054616; PMID:1385409
A;Accession: A44461
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <POW>
A;Cross-references: GB:M92301
A;Experimental source: skeletal muscle
A;Note: sequence extracted from NCBI backbone (NCBIP:118128); sequence inconsistent with
R;Iles, D.E.; Segers, B.; Segers, R.C.; Monsieus, K.; Heytens, L.; Halsall, P.J.; Hopk
Hum. Mol. Genet. 2, 863-868, 1993
A;Title: Genetic mapping of the beta 1- and gamma-subunits of the human skeletal muscle
thermia susceptibility.
A;Reference number: I38001; MUID:93372845; PMID:8395940
A;Accession: I38001
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 146-183 <ILE>
A;Cross-references: EMBL:Z21725; NID:g38562; PIDN:CAA79824.1; PID:g38563
A;Note: submitted to the EMBL Data Library, February 1993
C;Superfamily: human voltage-dependent calcium channel beta chain

Query Match 100.0%; Score 35; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 224 PPPHG 228

RESULT 29
T20482
hypothetical protein F01G10.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20482
R;Hembry, C.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19280
A;Accession: T20482
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-486 <WIL>
A;Cross-references: UNIPROT:O17768; EMBL:Z81055; PIDN:CAB02898.1; GSPDB:GN00022; CESP:F0
A;Experimental source: clone F01G10
C;Genetics:
A;Gene: CESP:F01G10.10

A;Map position: 4
A;Introns: 36/3; 73/2; 114/2; 154/1; 190/3; 211/3; 287/1; 364/3; 398/3; 449/2

Query Match 100.0%; Score 35; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 283 PPPHG 287

RESULT 30
JC6203
SP8 binding protein homolog - cucumber
C;Species: Cucumis sativus (cucumber)
C;Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C;Accession: JC6203
R;Kim, D.J.; Smith, S.M.; Leaver, C.J.
Gene 185, 265-269, 1997
A;Title: A cDNA encoding a putative SPF1-type DNA-binding protein from cucumber.
A;Reference number: JC6203; MUID:97208883; PMID:9055825
A;Accession: JC6203
A;Molecule type: mRNA
A;Residues: 1-509 <KIM>
A;Cross-references: UNIPROT:Q39658; GB:I44134; NID:g927024; PIDN:AAC37515.1; PID:g927025
C;Comment: This protein is a DNA-binding protein consisting of a monomer. It is involved
C;Genetics:
A;Gene: sz71
C;Keywords: DNA binding

Query Match 100.0%; Score 35; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 226 PPPHG 230

RESULT 31
T49422
RAD57 related protein [imported] - Neurospora crassa
N;Alternate names: protein B17C10.30
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49422
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49422
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <SCH>
A;Cross-references: UNIPROT:Q9P6E6; EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.30
A;Experimental source: BAC clone B17C10; strain OR74A
C;Genetics:
A;Gene: NCSP:B17C10.30
A;Map position: 6
A;Introns: 31/3

Query Match 100.0%; Score 35; DB 2; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 158 PPPHG 162

RESULT 32
S61920
B-alpha pheromone-receptor 2 - bracket fungus (Schizophyllum commune) (fragment)

C;Species: Schizophyllum commune
C;Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S61920
R;Wendland, J.; Vaillancourt, L.J.; Hegner, J.; Lengeler, K.B.; Laddison, K.J.; Specht, EMBO J. 14, 5271-5278, 1995
A;Title: The mating-type locus Balphal of Schizophyllum commune contains a pheromone receptor
A;Reference number: S61919; MUID:96080162; PMID:7489716
A;Accession: S61920
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-518 <WEN>
A;Cross-references: UNIPROT:Q05659; EMBL:X91168; NID:g1143553; PIDN:CAA62595.1; PID:g114
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C;Genetics:
A;Introns: 71/1; 108/3; 131/1; 177/2; 320/3

Query Match 100.0%; Score 35; DB 2; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 203 PPPHG 207

RESULT 33
I65767
L-type voltage-gated calcium channel B subunit - human
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 18-Aug-2000
C;Accession: I65767
R;Collin, T.; Wang, J.; Nargeot, J.; Schwartz, A.
Circ. Res. 72, 1337-1344, 1993
A;Title: Molecular cloning of three isoforms of the L-type voltage-dependent calcium channel
A;Reference number: I52859; MUID:93265672; PMID:7916667
A;Accession: I65767
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-522 <RES>
A;Cross-references: GB:L06112; NID:g187018; PIDN:AAA36169.1; PID:g187019
C;Superfamily: human voltage-dependent calcium channel beta chain

Query Match 100.0%; Score 35; DB 2; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 248 PPPHG 252

RESULT 34
A41347
calcium channel protein beta chain, dihydropyridine-sensitive, skeletal muscle - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
C;Accession: A41347
R;Ruth, P.; Roehrkasten, A.; Biel, M.; Bosse, E.; Regulla, S.; Meyer, H.E.; Flockerzi, V Science 245, 1115-1118, 1989
A;Title: Primary structure of the beta subunit of the DHP-sensitive calcium channel from
A;Reference number: A41347; MUID:89368946; PMID:2549640
A;Accession: A41347
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-524 <RUT>
A;Cross-references: UNIPROT:P19517; GB:M25817
C;Superfamily: human voltage-dependent calcium channel beta chain
C;Keywords: skeletal muscle

Query Match 100.0%; Score 35; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 249 PPPHG 253

RESULT 35
T00677
hypothetical protein At2g43970 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F6E13.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00677; G84872
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.
A;Reference number: Z14180
A;Accession: T00677
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-529 <ROU>
A;Cross-references: UNIPROT:O80567; EMBL:AC004005; NID:g3212846; PID:g3212854
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84872
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-529 <STO>
A;Cross-references: GB:AE002093; NID:g3212854; PIDN:AAC23405.1; GSPDB:GN00139
C;Genetics:
A;Gene: F6E13.10; At2g43970
A;Map position: 2
A;Introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 100.0%; Score 35; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
Db 76 PPPHG 80

RESULT 36
G84825
probable CCH-type zinc finger protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84825
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84825
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-597 <STO>
A;Cross-references: UNIPROT:Q9XEE6; GB:AE002093; NID:g6598933; PIDN:AAF18728.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g40140
A;Map position: 2

Query Match 100.0%; Score 35; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5

Db 456 PPPHG 460
|||||
RESULT 37
E86194
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E86194
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86194
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-614 <STO>
A;Cross-references: UNIPROT:Q9LNF0; GB:AE005172; NID:g9810458; PIDN:AAF80119.1; GSPDB:GN
C;Genetics:
A;Map position: 1
Query Match 100.0%; Score 35; DB 2; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
|||||
Db 46 PPPHG 50
RESULT 38
S48424
hypothetical protein YIL055C - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S48424
R;Smith, V.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48407
A;Accession: S48424
A;Molecule type: DNA
A;Residues: 1-627 <SMI>
A;Cross-references: UNIPROT:P40523; GB:Z47047; EMBL:Z38060; NID:g603997; PID:g763291; GS
C;Genetics:
A;Gene: MIPS:YIL055C
A;Cross-references: SGD:S0001317
A;Map position: 9L
Query Match 100.0%; Score 35; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
|||||
Db 222 PPPHG 226
RESULT 39
C87048
probable ABC transporter, ATP-binding component ML1113 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87048
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hd
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: C87048
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-629 <STO>
A;Cross-references: UNIPROT:Q9CC89; GB:AL450380; NID:g13093097; PIDN:CAC31494.1; GSPDB:GB
C;Genetics:
A;Gene: ML1113
C;Superfamily: Escherichia coli ABC transporter mdlA; ATP-binding cassette homology
Query Match 100.0%; Score 35; DB 2; Length 629;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
|||||
Db 382 PPPHG 386
RESULT 40
A49585
Na+ channel protein, amiloride-sensitive - human
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 09-Jul-2004
C;Accession: A49585; I51911
R;Voilley, N.; Lingueglia, E.; Champigny, G.; Mattei, M.G.; Waldmann, R.; Lazdunski, M.;
Proc. Natl. Acad. Sci. U.S.A. 91, 247-251, 1994
A;Title: The lung amiloride-sensitive Na+ channel: biophysical properties, pharmacology,
A;Reference number: A49585; MUID:94105144; PMID:8278374
A;Accession: A49585
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-669 <RES>
A;Cross-references: UNIPROT:P37088; EMBL:X76180; NID:g452649; PIDN:CAA53773.1; PID:g45264
R;McDonald, F.J.; Snyder, P.M.; McCray, P.B.
Am. J. Physiol. 266, L728-L734, 1994
A;Title: Cloning, expression, and tissue distribution of a human amiloride-sensitive Na+
A;Reference number: I51911; MUID:94295729; PMID:8023962
A;Accession: I51911
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-669 <RE2>
A;Cross-references: GB:L29007; NID:g493125; PIDN:AAA21813.1; PID:g493605
C;Genetics:
A;Map position: l2p13
C;Superfamily: human amiloride-sensitive sodium channel protein; fibronectin type I repea
Query Match 100.0%; Score 35; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
|||||
Db 195 PPPHG 199
Search completed: June 15, 2005, 14:22:51
Job time : 19.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-10
Perfect score: 35
Sequence: 1 PPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	35	100.0	33	2	O02832	002832 gorilla gor
2	35	100.0	53	2	Q9V8Y1	Q9v8y1 drosophila
3	35	100.0	62	2	Q6UB47	Q6ub47 hyaloperono
4	35	100.0	72	1	PN2D_LITSE	Q962a9 litopenaeus
5	35	100.0	80	2	Q9VPZ1	Q9vpz1 drosophila
6	35	100.0	93	2	Q94J14	Q94jl4 oryza sativ
7	35	100.0	94	2	Q62CG2	Q62cg2 burkholderi
8	35	100.0	94	2	Q63KL8	Q63kl8 burkholderi
9	35	100.0	108	2	Q655P3	Q655p3 oryza sativ
10	35	100.0	111	2	Q9YCY2	Q9ycy2 aeropyrum p
11	35	100.0	119	2	Q67867	Q67867 hepatitis b
12	35	100.0	134	2	Q7YWR5	Q7ywr5 caenorhabdi
13	35	100.0	135	2	Q7VVE3	Q7vve3 bordetella
14	35	100.0	135	2	Q7WBB4	Q7wbb4 bordetella
15	35	100.0	135	2	Q7WMT3	Q7wmt3 bordetella
16	35	100.0	147	1	SMR1_MOUSE	Q61900 mus musculu
17	35	100.0	151	2	Q8N1P8	Q8nlp8 homo sapien
18	35	100.0	152	2	Q8ABU0	Q8abu0 bacteroides
19	35	100.0	164	2	Q8GRK5	Q8grk5 oryza sativ
20	35	100.0	164	2	Q8GSL3	Q8gsl3 oryza sativ
21	35	100.0	164	2	Q657L2	Q657l2 oryza sativ
22	35	100.0	164	2	Q62HZ3	Q62hz3 burkholderi
23	35	100.0	164	2	Q63RT2	Q63rt2 burkholderi
24	35	100.0	165	2	Q6ZW55	Q6zw55 homo sapien
25	35	100.0	168	2	O06065	O06065 mycobacteri
26	35	100.0	169	2	O07429	O07429 mycobacteri
27	35	100.0	169	2	Q7U2M4	Q7u2m4 mycobacteri
28	35	100.0	170	2	Q07610	Q07610 rattus norv
29	35	100.0	173	2	O62063	O62063 caenorhabdi
30	35	100.0	173	2	O62789	O62789 sus scrofa
31	35	100.0	173	2	Q8LLM4	Q8llm4 triticum ae

32	35	100.0	173	2	Q9LF59	Q9lf59 arabidopsis
33	35	100.0	175	2	Q52724	Q52724 rhodococcus
34	35	100.0	176	2	Q8LEP5	Q8lep5 arabidopsis
35	35	100.0	176	2	Q9C4Z8	Q9c4z8 arabidopsis
36	35	100.0	181	2	Q28261	Q28261 canis famil
37	35	100.0	185	2	Q8ZY20	Q8zy20 pyrobaculum
38	35	100.0	192	2	Q67U43	Q67u43 oryza sativ
39	35	100.0	201	2	Q8ND66	Q8nd66 homo sapien
40	35	100.0	203	2	Q67IV2	Q67iv2 oryza sativ
41	35	100.0	206	1	PRP3_RAT	P04474 rattus norv
42	35	100.0	215	2	Q9YEL2	Q9yel2 aeropyrum p
43	35	100.0	216	2	Q9FYN6	Q9fyn6 oryza sativ
44	35	100.0	221	2	Q8RYZ6	Q8ryz6 oryza sativ
45	35	100.0	224	2	Q9A825	Q9a825 caulobacter

ALIGNMENTS

RESULT 1
O02832
ID O02832 PRELIMINARY; PRT; 33 AA.
AC O02832;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Huntingtin protein (Fragment).
GN Name=IT15;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326790; PubMed=8766138;
RA Pecheux C., Gall A.L., Kaplan J.C., Dode C.;
RT "Sequence analysis of the CAG triplet repeats region in the Huntington
disease gene (IT15) in several mammalian species.";
RL Ann. Genet. 39:81-86(1996).
DR EMBL; S83377; AAB50771.1; -.
FT NON TER 33
SQ SEQUENCE 33 AA; 3393 MW; 44F558CE5636722A CRC64;

Query Match 100.0%; Score 35; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
5 PPPHG 9

RESULT 2
Q9V8Y1
ID Q9V8Y1 PRELIMINARY; PRT; 53 AA.
AC Q9V8Y1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG15126-PA.
GN ORFNames=CG15126;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]

RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003795; AAF57523.1; -.
DR IntAct; Q9V8Y1; -.
DR FlyBase; FBgn0040729; CG15126.
SQ SEQUENCE 53 AA; 5388 MW; 68828750C9C82142 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
2 PPPHG 6

RESULT 3
Q6UB47 PRELIMINARY; PRT; 62 AA.
ID Q6UB47
AC Q6UB47;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cysteine rich (Fragment).
OS Hyaloperonospora parasitica.
OC Eukaryota; stramenopiles; Oomycetes; Peronosporales; Peronosporaceae;
OC Hyaloperonospora.
OX NCBI_TaxID=123356;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maks9;
RA Bittner-Eddy P., Allen R., Rehmany A., Birch P., Beynon J.;
RT "Use of suppression subtractive hybridization to identify downy mildew
RL genes expressed during infection of Arabidopsis thaliana.";
RL Mol. Plant Pathol. 0:0-0(2003).
DR EMBL; AY373968; AAQ83520.1; -.
FT NON TER 62
SQ SEQUENCE 62 AA; 6729 MW; D883FC9E4E838A58 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db |||||
31 PPPHG 35

RESULT 4
PN2D_LITSE STANDARD; PRT; 72 AA.
ID PN2D_LITSE
AC Q962A9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Penaeidin-2d precursor (Pen-2d).
OS Litopenaeus setiferus (Atlantic white shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Litopenaeus.
OX NCBI_TaxID=64468;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hemocyte;
RX MEDLINE=22226701; PubMed=12242595; DOI=10.1007/s00251-002-0487-z;
RA Cuthbertson B.J., Shepard E.F., Chapman R.W., Gross P.S.;
RT "Diversity of the penaeidin antimicrobial peptides in two shrimp
RT species.";
RL Immunogenetics 54:442-445(2002).
CC -!- FUNCTION: Antibacterial and antifungal activity. Presents chitin-
CC binding activity (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic granules of hemocytes and to a

CC lesser extent in small granules of hemocytes.
CC -!- SIMILARITY: Belongs to the penaeidin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AY039205; AAK83453.1; -.
DR HSSP; P81058; 1UEO.
DR InterPro; IPR009226; Penaeidin.
DR Pfam; PF05927; Penaeidin; 1.
KW Amidation; Antibiotic; Chitin-binding; Fungicide;
KW Pyrrolidone carboxylic acid; Signal.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 71 Penaeidin-2d.
FT DOMAIN 29 43 Pro-rich.
FT DISULFID 45 59 By similarity.
FT DISULFID 48 66 By similarity.
FT DISULFID 60 67 By similarity.
FT MOD_RES 22 22 Pyrrolidone carboxylic acid (By
FT similarity).
FT MOD_RES 71 71 Lysine amide (G-72 provides amide group)
FT (By similarity).
FT SEQUENCE 72 AA; 7717 MW; 7DCE24C8F3D3F182 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 33 PPPHG 37

RESULT 5
Q9VPZ1
ID Q9VPZ1 PRELIMINARY; PRT; 80 AA.
AC Q9VPZ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG5011-PA.
GN ORFNames=CG5011;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003587; AAF51393.1; -.
DR IntAct; Q9VPZ1; -.
DR FlyBase; FBgn0040723; CG5011.
SQ SEQUENCE 80 AA; 8699 MW; D57F2D6B3443B4BF CRC64;

Query Match 100.0%; Score 35; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
|||||

Db 12 PPPHG 16

RESULT 6

Q94J14

ID Q94J14 PRELIMINARY; PRT; 93 AA.

AC Q94J14;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE P0481E12.34 protein.

GN Name=P0481E12.34;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=12447438; DOI=10.1038/nature01184;

RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,

RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,

RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,

RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,

RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,

RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,

RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,

RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,

RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,

RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,

RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,

RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,

RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,

RA Yano M., Jiang J., Gojobori T.;

RT "The genome sequence and structure of rice chromosome 1.";

RL Nature 420:312-316(2002).

DR EMBL; AP003076; BAB56049.1; --

DR Gramene; Q94J14; --

SQ SEQUENCE 93 AA; 9989 MW; 6EDA478C67F82098 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 85 PPPHG 89

RESULT 7

Q62CG2

ID Q62CG2 PRELIMINARY; PRT; 94 AA.

AC Q62CG2;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE DNA-binding protein.

GN ORFNames=BMAA0931;

OS Burkholderia mallei ATCC 23344.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Burkholderiaceae; Burkholderia.

OX NCBI_TaxID=243160;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 23344;

RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,

RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,

RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,

RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,

RA Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,

RA Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,

RA Zhou L., Fraser C.M.;

RT "Structural flexibility in the Burkholderia mallei genome.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).

DR EMBL; CP000011; AAU46112.1; --

KW DNA-binding.

SQ SEQUENCE 94 AA; 9817 MW; E0DA5022E10EE8FF CRC64;

Query Match 100.0%; Score 35; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 75 PPPHG 79

RESULT 8

Q63KL8

ID Q63KL8 PRELIMINARY; PRT; 94 AA.

AC Q63KL8;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Putative DNA-binding protein.

GN ORFNames=BPSS1343;

OS Burkholderia pseudomallei K96243.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Burkholderiaceae; Burkholderia.

OX NCBI_TaxID=272560;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K96243;

RX PubMed=15377794;

RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,

RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,

RA Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.R.,

RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,

RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,

RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,

RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,

RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,

RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,

RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.,

RT "Genomic plasticity of the causative agent of melioidosis,

RT Burkholderia pseudomallei.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).

DR EMBL; BX571966; CAH3814.1; --

KW DNA-binding.

SQ SEQUENCE 94 AA; 9817 MW; E0DA5022E10EE8FF CRC64;

Query Match 100.0%; Score 35; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 75 PPPHG 79

RESULT 9

Q655P3

ID Q655P3 PRELIMINARY; PRT; 108 AA.

AC Q655P3;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein B1011A07.27.

GN Name=B1011A07.27;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.


```

RESULT 13
Q7VVE3
ID Q7VVE3 PRELIMINARY; PRT; 135 AA.
AC Q7VVE3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome ubiquinol oxidase subunit IV.
GN OrderedLocusNames=BP2728;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640419; CAE43005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005171; COX4_pro.
DR Pfam; PF03626; COX4_pro; 1.
DR Complete proteome.
KW SEQUENCE 135 AA; 14848 MW; EF2B611E6F958A5D CRC64;
SQ
Query Match 100.0%; Score 35; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 12 PPPHG 16

RESULT 14
Q7WBB4
ID Q7WBB4 PRELIMINARY; PRT; 135 AA.
AC Q7WBB4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome ubiquinol oxidase subunit IV.
GN OrderedLocusNames=BPPI091;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640419; CAE43005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005171; COX4_pro.
DR Pfam; PF03626; COX4_pro; 1.
DR Complete proteome.
KW SEQUENCE 135 AA; 14848 MW; EF2B611E6F958A5D CRC64;
SQ
Query Match 100.0%; Score 35; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 12 PPPHG 16

RESULT 15
Q7WMT3
ID Q7WMT3 PRELIMINARY; PRT; 135 AA.
AC Q7WMT3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome ubiquinol oxidase subunit IV.
GN OrderedLocusNames=BB1307;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640441; CAE31805.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005171; COX4_pro.
DR Pfam; PF03626; COX4_pro; 1.
DR Complete proteome.
KW SEQUENCE 135 AA; 14848 MW; EF2B611E6F958A5D CRC64;
SQ
Query Match 100.0%; Score 35; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 12 PPPHG 16

RESULT 16
SMR1_MOUSE
ID SMR1_MOUSE STANDARD; PRT; 147 AA.
AC Q61900;
```

```

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640426; CAE36392.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005171; COX4_pro.
DR Pfam; PF03626; COX4_pro; 1.
DR Complete proteome.
KW SEQUENCE 135 AA; 14848 MW; EF2B611E6F958A5D CRC64;
SQ
Query Match 100.0%; Score 35; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 12 PPPHG 16

RESULT 17
Q7WWT3
ID Q7WWT3 PRELIMINARY; PRT; 135 AA.
AC Q7WWT3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome ubiquinol oxidase subunit IV.
GN OrderedLocusNames=BB1307;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640441; CAE31805.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005171; COX4_pro.
DR Pfam; PF03626; COX4_pro; 1.
DR Complete proteome.
KW SEQUENCE 135 AA; 14848 MW; EF2B611E6F958A5D CRC64;
SQ
Query Match 100.0%; Score 35; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 12 PPPHG 16

RESULT 18
SMR1_MOUSE
ID SMR1_MOUSE STANDARD; PRT; 147 AA.
AC Q61900;
```

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Submaxillary gland androgen regulated protein 1 precursor (Salivary protein MSG1).

DE Name=Smr1; Synonyms=MSG1;

GN Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Submaxillary gland;

RX MEDLINE=94252564; PubMed=8194749; DOI=10.1016/0378-1119(94)90258-5;

RA Tronik-Le Roux D., Senorale-Pose M., Rougeon F.;

RT "Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice show extensive evolutionary divergence in the protein coding region.";

RL Gene 142:175-182(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Salivary gland;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: May play a role in protection or detoxification.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Secreted into saliva by submaxillary gland.

CC -!- SIMILARITY: Belongs to the PROL1/PROL3 family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X71629; CAA50636.1; -.

DR EMBL; BC031806; AAH31806.1; -.

DR EMBL; BC031921; AAH31921.1; -.

DR EMBL; BC034553; AAH34553.1; -.

DR EMBL; BC055857; AAH55857.1; -.

DR PIR; I48669; S37485.

DR MGD; MGI:102763; Smr1.

KW Repeat; Signal.

FT SIGNAL 1 22 Potential.

FT CHAIN 23 147 Submaxillary gland androgen regulated protein 1.

FT FT

FT DOMAIN 43 78 3 X 12 AA tandem repeats of G-P-G-I-G-R-P-[HP]-P-P-P-[PF].

FT REPEAT 43 54 1.

FT REPEAT 55 66 2.

FT REPEAT 67 78 3.

FT DOMAIN 73 84 Poly-Pro.

SQ SEQUENCE 147 AA; 15544 MW; 07D64E22F1BF58C2 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 147;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 39 PPPHG 43

RESULT 17

Q8N1P8

ID Q8N1P8 PRELIMINARY; PRT; 151 AA.

AC Q8N1P8;

DT 01-OCT-2002 (TReMBLrel. 22, Created)

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)

DE Hypothetical protein FLJ38012.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Tongue;

RX PubMed=14702039; DOI=10.1038/ng1285;

RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;

RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.";

RL Nat. Genet. 36:40-45(2004).

DR EMBL; AK095331; BAC04532.1; -.

SQ SEQUENCE 151 AA; 16189 MW; 091EADD64FCEEE89 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 151;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 19 PPPHG 23

RESULT 18

Q8ABU0

ID Q8ABU0 PRELIMINARY; PRT; 152 AA.

AC Q8ABU0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT0020;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."
RL Science 299:2074-2076(2003).
DR EMBL; AE016926; AAO75127.1; --.
KW Complete proteome.
SQ SEQUENCE 152 AA; 17665 MW; 1991F9218509F8D9 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 143 PPPHG 147

RESULT 19
Q8GRK5
ID Q8GRK5 PRELIMINARY; PRT; 164 AA.
AC Q8GRK5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ112_E08.108 (Hypothetical protein
OJ1112_E08.111) (Hypothetical protein OJ1112_E08.113) (Hypothetical
protein OJ1720_F04.121) (Hypothetical protein OJ1720_F04.124).
GN Name=OJ1112_E08.108;
GN Synonyms=OJ1112_E08.111, OJ1112_E08.113, OJ1720_F04.121,
OJ1720_F04.124;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
clone:OJ1112_E08.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
clone:OJ1720_F04.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP003705; BAC15809.1; --.
DR EMBL; AP005296; BAD31279.1; --.
DR Gramene; Q8GRK5; --.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR008994; Nucleic_acid_OB.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17805 MW; E6F13AF5FA7E4E95 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 164;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 49 PPPHG 53

RESULT 20
Q8GSL3
ID Q8GSL3 PRELIMINARY; PRT; 164 AA.
AC Q8GSL3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ1112_E08.110 (Hypothetical protein
OJ1720_F04.123).
GN Name=OJ1112_E08.110; Synonyms=OJ1720_F04.123;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
clone:OJ1112_E08.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
clone:OJ1720_F04.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP003705; BAC15808.1; --.
DR EMBL; AP005296; BAD31279.1; --.
DR Gramene; Q8GSL3; --.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR008994; Nucleic_acid_OB.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17708 MW; 02369EB26D9ADB1C CRC64;

Query Match 100.0%; Score 35; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 49 PPPHG 53

RESULT 21
Q657L2
ID Q657L2 PRELIMINARY; PRT; 164 AA.
AC Q657L2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein P0468H06.16;
GN Name=P0468H06.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,

RA	Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,	
RA	Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,	
RA	Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,	
RA	Machita K., Machara T., Mizuno H., Mizubayashi T., Mukai Y.,	
RA	Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,	
RA	Namiki N., Negishi M., Ohta I., Ono N., Saij S., Sakai K., Shibata M.,	
RA	Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,	
RA	waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,	
RA	Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,	
RA	Yano M., Jiang J., Gojobori T.;	
RT	"The genome sequence and structure of rice chromosome 1.";	
RL	Nature 420:312-316(2002).	
DR	EMBL; AF003075; BAD45005.1; -.	
KW	Hypothetical protein.	
SQ	SEQUENCE 164 AA; 18134 MW; 6CF81BD88C3C30C7 CRC64;	
Qy	1 PPPHG 5	
Db	135 PPPHG 139	
RESULT 22		
Q62HZ3		
ID	Q62HZ3 PRELIMINARY; PRT; 164 AA.	
AC	Q62HZ3;	
DT	25-OCT-2004 (TrEMBLrel. 28, Created)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)	
DE	TonB domain protein.	
GN	ORFNames=BMA2105;	
OS	Burkholderia mallei ATCC 23344.	
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;	
OC	Burkholderiaceae; Burkholderia.	
OX	NCBI_TaxID=243160;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 23344;	
RA	Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,	
RA	Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,	
RA	Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,	
RA	Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,	
RA	Mohammoud Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,	
RA	Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,	
RA	Zhou L., Fraser C.M.;	
RT	"Structural flexibility in the Burkholderia mallei genome.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).	
DR	EMBL; CP00010; AAU50017.1; -.	
SQ	SEQUENCE 164 AA; 17543 MW; 3B29EF2562073ABB CRC64;	
Query Match	100.0%; Score 35; DB 2; Length 164;	
Best Local Similarity	100.0%; Pred. No. 2.3e+02;	
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5	
Db	34 PPPHG 38	
RESULT 23		
Q63RT2		
ID	Q63RT2 PRELIMINARY; PRT; 164 AA.	
AC	Q63RT2;	
DT	25-OCT-2004 (TrEMBLrel. 28, Created)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)	
DE	Putative lipoprotein.	
GN	ORFNames=BPSL2589;	
OS	Burkholderia pseudomallei K96243.	
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;	

OC	Burkholderiaceae; Burkholderia.	
OX	NCBI_TaxID=272560;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=K96243;	
RX	PubMed=15377794;	
RA	Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,	
RA	Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,	
RA	Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,	
RA	Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,	
RA	Chillingworth T., Cronin A., Crosset B., Davis P., DeShazer D.,	
RA	Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,	
RA	Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,	
RA	Rabbinowitsch E., Rutherford K., Sanders M., Simmonds M.,	
RA	Songsivilai S., Stevens K., Tumapa S., Vesaratchavest M.,	
RA	Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;	
RT	"Genomic plasticity of the causative agent of melioidosis,	
RT	Burkholderia pseudomallei.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).	
DR	EMBL; BX571965; CAH36597.1; -.	
KW	Lipoprotein.	
SQ	SEQUENCE 164 AA; 17559 MW; 3B360F257DE73ABB CRC64;	
Query Match	100.0%; Score 35; DB 2; Length 164;	
Best Local Similarity	100.0%; Pred. No. 2.3e+02;	
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5	
Db	34 PPPHG 38	
RESULT 24		
Q6ZW55		
ID	Q6ZW55 PRELIMINARY; PRT; 165 AA.	
AC	Q6ZW55;	
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Hypothetical protein FLJ41588.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Tongue;	
RA	Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,	
RA	Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,	
RA	Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,	
RA	Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,	
RA	Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,	
RA	Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,	
RA	Masuko Y., Nagai K., Isogai T.;	
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AK123582; BAC85651.1; -.	
SQ	SEQUENCE 165 AA; 17648 MW; 9CADD2C8FDFCB90F CRC64;	
Query Match	100.0%; Score 35; DB 2; Length 165;	
Best Local Similarity	100.0%; Pred. No. 2.3e+02;	
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 PPPHG 5	
Db	33 PPPHG 37	
RESULT 25		
O06065		
ID	O06065 PRELIMINARY; PRT; 168 AA.	
AC	O06065;	
DT	01-JUL-1997 (TrEMBLrel. 04, Created)	
DT	01-JUL-1997 (TrEMBLrel. 04, Last sequence update)	

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein MLCC1622.02 (Hypothetical protein ML2605).
GN Name=MLCC1622.02; OrderedLocusNames=ML2605;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93188700; PubMed=8446027;
RA Eiglmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT "Use of an ordered cosmid library to deduce the genomic organization
of Mycobacterium leprae.";
RL Mol. Microbiol. 7:197-206(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Hamlin N., Churcher C.M.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL; 295398; CAB08794.1; --.
DR EMBL; AL583926; CAC32137.1; --.
DR PIR; C87235; C87235.
DR Leproma; ML2605; --.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 168 AA; 18602 MW; 6A6ED74C046359B1 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 131 PPHG 135

RESULT 26
O07429 PRELIMINARY; PRT; 169 AA.
ID O07429 Q7DAB3;
AC O07429; Q7DAB3;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MT0194, Rv0185;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,

RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; BX842572; CAB09736.1; --.
DR EMBL; AE000516; AAK44414.1; --.
DR PIR; D70906; D70906.
DR TIGR; MT0194; --.
DR TuberculList; Rv0185; --.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 169 AA; 18397 MW; 8E7677B0190054F3 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 132 PPHG 136

RESULT 27
Q7U2M4 PRELIMINARY; PRT; 169 AA.
ID Q7U2M4 Q7U2M4;
AC Q7U2M4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Mb0191.
GN OrderedLocusNames=Mb0191;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248334; CAD93055.1; --.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 169 AA; 18397 MW; 8E7677B0190054F3 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 132 PPPHG 136

RESULT 28

Q07610
ID Q07610 PRELIMINARY; PRT; 170 AA.
AC Q07610;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proline-rich proteoglycan.
GN Name=PRPG1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Parotid;
RX MEDLINE=93388626; PubMed=8376404;
RA Castle A.M., Castle J.D.;
RT "Novel secretory proline-rich proteoglycans from rat parotid. Cloning
RT and characterization by expression in Atr-20 cells.";
RL J. Biol. Chem. 268:20490-20496(1993).
DR EMBL; L17317; AAA03073.1; -.
DR PIR; A48013; A48013.
SQ SEQUENCE 170 AA; 17840 MW; 85AF34C40B45AA03 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 170;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 80 PPPHG 84

RESULT 29

O62063
ID O62063 PRELIMINARY; PRT; 173 AA.
AC O62063;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein C16D6.1.
GN ORFNames=C16D6.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81472; CAB03889.1; -.
DR PIR; T19341; T19341.
DR WormBase; WBGene00007634; C16D6.1.
DR WormPep; C16D6.1; CE08230.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19001 MW; 16D04C1BBDF76265 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 72 PPPHG 76

RESULT 30

O62789
ID O62789 PRELIMINARY; PRT; 173 AA.
AC O62789;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Collagen VIII (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic smooth muscle;
RA Reichenberg S., Plenz G., Robenek H.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054891; AAC08999.1; -.
DR HSSP; Q00780; I091.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Collagen.
FT NON_TER 1 1
FT NON_TER 173 173
SQ SEQUENCE 173 AA; 18973 MW; 03489B6FBAAB6CAD CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 7 PPPHG 11

RESULT 31

O8LLM4
ID O8LLM4 PRELIMINARY; PRT; 173 AA.
AC O8LLM4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Thaumatin-like protein 1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Vroh Bi I., Kolb F., Boze L., Domier L.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389883; AAM69454.1; -.
DR HSSP; P33679; 1DU5.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; Thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD001321; Thaumatin; 1.
DR PROSITE; PS00316; THAUMATIN; 1.

SQ SEQUENCE 173 AA; 17646 MW; EC172D12ACD944D3 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 145 PPPHG 149

RESULT 32

Q9LF59 ID Q9LF59 PRELIMINARY; PRT; 173 AA.

AC Q9LF59;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Glycine/proline-rich protein (At5g17650).

GN Name=K10A8_130;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,

RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;

RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Shinn P., Carninci P., Hayashizaki Y.,

RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,

RA Seki M., Shinozaki K., Ecker J.R.;

RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AL391151; CAC01909.1; -.

DR EMBL; BT012163; AAS76258.1; -.

DR PIR; T51469; T51469.

SQ SEQUENCE 173 AA; 18536 MW; E3510947AA98BC0A CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 49 PPPHG 53

RESULT 33

Q52724 ID Q52724 PRELIMINARY; PRT; 175 AA.

AC Q52724;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Rhodococcus corallinus (NRRL 15444B) N-ethylammeline chlorohydrolase (trzA) (Fragment).

OS Rhodococcus corallinus.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Gordoniaceae; Gordonia.

OX NCBI_TaxID=36822;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96011356; PubMed=7592318;

RA Shao Z.Q., Seffens W., Mulbry W., Behki R.M.;

RT "Cloning and expression of the s-triazine hydrolase gene (trzA) from Rhodococcus corallinus and development of Rhodococcus recombinant

RT strains capable of dealkylating and dechlorinating the herbicide atrazine.";

RL J. Bacteriol. 177:5748-5755(1995).

DR EMBL; L16534; AAA90930.1; -.

DR PIR; T46665; T46665.

DR HSSP; Q12178; 1P6O.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

KW Hydrolase.

FT NON_TER 1 1

SQ SEQUENCE 175 AA; 19406 MW; 881A0E9059A62528 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 129 PPPHG 133

RESULT 34

Q8LEP5 ID Q8LEP5 PRELIMINARY; PRT; 176 AA.

AC Q8LEP5;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Putative glycine and proline-rich protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22088475; PubMed=12093376;

RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,

RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;

RT "Full-length messenger RNA sequences greatly improve genome annotation.";

RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).

RN [2]

RP SEQUENCE FROM N.A.

RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,

RA Feldmann K.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AY085309; AAM62540.1; -.

DR InterPro; IPR006031; YPPPX.

DR Pfam; PF02162; YPPPX; 8.

SQ SEQUENCE 176 AA; 18434 MW; E84E78E58D435547 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||

Db 50 PPPHG 54

RESULT 35

Q9C4Z8 ID Q9C4Z8 PRELIMINARY; PRT; 176 AA.

AC Q9C4Z8;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein F27M3_5 (Hypothetical protein F5M6.24).

GN Name=F27M3_5; Synonyms=F5M6.24;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ansoerge W., Wirkner U., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL834172; CAD38870.1; --
DR HSSP; P42771; 2ASE.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 2.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 2.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 201 AA; 21709 MW; 1E26E07B168873E5 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 188 PPPHG 192

RESULT 40
Q67IV2
ID Q67IV2 PRELIMINARY; PRT; 203 AA.
AC Q67IV2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNBa0048L03.37.
GN Name=OSJNBa0048L03.37;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Fujisawa M.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OSJNBa0048L03."
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP007205; BAD38589.1; --
KW Hypothetical protein.
SQ SEQUENCE 203 AA; 21805 MW; 204D6EED4120FD35 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 9 PPPHG 13